

FIGURE 1

CCAGGTCCAAC TGACCTCGGTTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGCCTCAGGAGGTGCCCTCAGGCCAGTGGCCTGAGGCCAGC
AAGGGTAGGGTCCATCTCAGTCCCAGGACACAGCAGGCCACCAGGCCACGCCCTGGC
TCCAGCAGCATCAGCAGCCCCCAGGACGGGGGAGGCACAGGTGGCCCCACCACCCGGAGG
AGCAGCTCTGCCCCGTCCGGGG**ATG**ACTGATTCTCTCCGCCAGGCCACCCAGAGGAGA
AGGCCACCCGCCCTGGAGGCACAGGCCATGAGGGCTCTCAGGAGGTGCTGATGTGGCT
TCTGGTGTGGCAGTGGGGCACAGAGCACGCCCTACGGGCCGGCCGTTAGGGTGTGCT
GTCCCCGGCTCACGGGACCCGTCTCCAGTCGTTCTGCAAGCGTGTGATACGCCCTTC
TCACCACCTGCGACGGCACCGGCCCTGCAAGCACCTACCGAACATTATAGGACGCCCTAC
CGCCGAGCCCTGGCTGGGCCCTGCCAGGCCCTGCAAGCGTGTGCTGCCCGCTGGAAGAG
GACCAGCGGGCTCTGGCTGGGCCCTGTGGAGCAGCAATATGCCAGGCCCATGCCGAACGGAG
GGAGCTGTGTCAGCCCTGGGCCCTGCCAGGCCCTGCAAGGATGGCGGGTGACACTTGCAG
TCAGATGTGGATGAATGCACTGAGTCTAGGAGGGCGGCTGTCCCCAGCGCTGCATCAACACCGC
CGGCAGTTACTGGTGCAGTGTTGGAGGGCACAGCCTGCTGCAGACGGTACACTCTG
TGCCCAAGGGAGGGCCCCCAGGGTGGCCCCAACCCGACAGGAGTGGACAGTGCAATGAAG
GAAGAAGTGCAAGAGGCTGCACTGCAAGGGTGGACCTGCTGGAGGAGAAAGCTGCACTGGTGC
GGCCCCACTGCAAGCCTGCCCTGCAGGCACTGGAGCATGGCTCCGGACCCGGCAGCC
TCCTGGTGCACTCCTCCAGCAGCTGGCCGCATCGACTCCCTGAGCGAGCAGATTCCCTC
CTGGAGGAGCAGCTGGGCTCTGCTCCCTGCAAGAAAGACTCG**TG**ACTGCCAGGCCAGG
CTGGACTGAGCCCCCTCACGCCCTGCAGCCCCATGCCCTGCCAACATGCTGGGGTC
CAGAAGCCACCTGGGGTCACTGAGCGGAAGGCCAGGCAGGGCTTCTCTTTCTCCTC
CCCTTCCCTGGAGGGTCCCCAGACCCCTGGCATGGATGGCTGGATTTTTTGTGAAT
CCACCCCTGGCTACCCACCCCTGGTACCCCAACGGCATCCCAAGGCCAGGTGGCCCTCA
GCTGAGGGAGGTACGAGTCCCCCTGCTGGAGCCATGGCACAGGCCAGGCAGCC
CGGAGGCTGGGTGGGCCCTCAGTGGGGCTGCTGCCCTGACCCCCAGCACAAATAAATGAAA
CGTGAaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaAGGGCGGCCGACTCT
AGAGTCGACCTGCAGAAGCTGGCCGCATGCCCAACTGTTATTGCAGCTTATAATGGT
TACAAAT

FIGURE 2

MTDSPPPGHPEEKATPPGGTGHEGLSGGAADVVASGVGSGRHRARLPARPLGCVLSRAHDPV
SESFVQRVYQPFLTTCDGHACSTYRTIYTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC
GAAICQPPCRNGSCVQPGRCRCPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCW
EGHSLSAADGTLCKVPKGGRPVAPNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLAS
QALEHGLPDGSLLVHSFQQQLRIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 93-97, 270-274

N-myristoylation sites.

amino acids 19-25, 78-84, 97-103, 100-106, 103-109, 157-163,
191-197, 265-271

Amidation site.

amino acids 26-30

Aspartic acid and asparagine hydroxylation site.

amino acids 152-164

Cell attachment sequence.

amino acids 130-133

EGF-like domain cysteine pattern signature.

amino acids 123-135

FIGURE 3

CGCTCGCCCCGTCGCCCTCGCTCCCCGCAGAGTCCCCCGCGCAGCAGATGTGTGTGGG
GTCAGCCCCAGGGGGACTATGGTGAATTCCCGGCCCTACGCACTACTGGCCCTGATC
CGGTTCTGGTGCCTCGGCATACCAACATAGCCATCGACTTCGGGGAGCAGGCCCTGAA
CCGGGGCATGGTGTGCAAGGGAGATCGAGTCAGAGTGTGGCCAGCTACGGGCTGGCT
ACTCTGGTCAATGGTGTGCAAGGGAGATCGAGTCAGAGTGTGGCCAGCTACGGGCTGGCT
GTGAAACAGCAAGAGAGACAGGACCAAAGCGTCTGTATGGTGTGGCAGGGCCATCGC
TGCGCTCTTCACACACTGATAGCTTATAGTGTAGTTAGGATACTACATTATCAATAACTGC
ACCATGTGGACGAGTGGGGGGAGCAAGCAGAAGGGCCCTCTGTACCTCGCCGCCCTT
CTTTCATGGACGCAATGGCATGGGCCATCTGGCATTCTCTTAAACACAATAAGCTT
CTTGGTGGGATGTGCTCAATCTAGATGTCAAGCTCAGGGTGTGTTGTGACCTTTGC
TTCACAGTCACCTGGAAATGCCGGAGGCCCTGCTCATCCGATCTCTCCCTGTACATGGG
GCACTTGTGCGCTGCACCCCTGTGCTGGGCTACTACAAGAACATTCACGGCCATCATCCC
TGACAGAACTGGCCCGAGCTGGGGGGAGATGCAACAAATAAGAAAGATGCTGAGCTTGTG
GGCCTTGGCTCTAATTCTGCCACACAGAGAATCAGTCGGCTATTGTCAACCTCTTGT
TCCGGGGACCTTGTGGCAGTCTGCAACAGGAGCAGAGGAGTGGGAGATTGACGGCACACATA
CCCTGTGGGTACATGCCATACGGCTGGTGAAGGAAACTCGTGTGCTGTATCTGCTTGT
ACAAGAATAACCCAGCAACAAACTGTGAGCACGAGCAACAGTCACGGCAGCCACATC
AAGAAGTTACCTTCGTCGTGATGGAAATCTGATGACATCATGGAGTGGACTTGTGAGTTGGAC
ACCATGCTGTCGTGAGAAAATCTGATGACATCATGGAGTGGACTTGTGAGTTGGAC
TCTGTGTTGCTCTTGGGAGATCTCTCTCTTCCCGACTTCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAGAAAACCTTCGTCCTTGCCCCAGCTCTGCTGCG
GATCATGCTCTCTCATGCCAGCTCTGCTACCCCTACCTGGGGTGCACGGTGCAGCC
TGGCGTGGGCTCCCTCTGGGGCTTTGTGGGAAGATCACCAGTGTGCCATCGCTGCG
TGCTATGTCACCGAACAGGAGAAAAGAGATGGGAAGATGAGTCGGGACAGGGAGGGGAAAGA
CTCTGCCATGACAGAACATGCTCCGACAGAGGGAGTCAGACACATCTGGAAATGAGAGG
AGAATGAAATAAGGACAGGGACGGCATGGCACTGGCAGGGCAGGTCACTGGATGACACTTC
GGCATCATCTCTCCCTCTCCATGCTATTGTTCTCTTGTGTTGTTGTTGTTGTAAT
GAAAGAGGCCCTGATTTAAAGGTTCTGTCATTCTAGCATACTGGGTATGCTCACACT
GACGGGGGGGACCTGTAGTGTGTTCTTACTGTTGCTATGAAAAACAAACGAAACAACTGAC
TTCATACCCCTGCTCAGGAAACCAAAAGACACAGCTGCTCACGTTGACGTGTTGTCC
TCCCTCCCTGACAAATCTCTCTTGGAACCAAGGACTGCACTGTCGTCCTCGCCATCGCCCTGGT
CACCCCTGCAACAGCAGGCCAGAACATCTCTGTCCTCCCTCATGCTCTTAAGAATCACAGG
TTAAACACTGGCTCTTGTGTTGCTTCTTGGACTACATGGCCATCAAGAGATGGAGGCC
CGGTGCTCTTAAATTCTCTCTGCCACGGAGTCTGCAACACCATTACTCCACATGCA
GAGGGGGTGGCACGCTGCAAGCCGGAGTCCCCGTTCACACTGAGGAACGGAGACTGTGAC
CACAGCAGGGTCAAGAGTGGACAGAACATCTCCCTGGAGAAAGGGTTGGGAAATGCCCGGG
GGCAGCAAACACTGACATGGTGAATGATGACATTCTACTGCTGTTCTCTAGATCTGAGCAA
GCTGTCAGITCTCACCCCCACCGTGTATATACATGAGCTAACCTTTAAATTGTCAACAAA
GCGCATCTCACGATTCCAGACCCCTGCCATGACTTTCTGAAGGGCTGCTTTCCCTCGC
CTTCTGTCAGGTCGATTAGAGCAGTCAGATGGAGCAGTCACTGGAGCATCTTAACCTTG
TACAGTGAACACTGAGCTTAAAGTCTCATCCAGCATTCTAATGCCAGGTTGCTGTAGGGTAAC
TTGAGAATGAGATATATTACCTGGTTCTGCTATCTTAACTGATAACTCTGCGGTACAGGAA
ATACAGAGTCACTATGTTGCTCTCTGGAAATAACGCAATTGCAAACTGAGACATTTATAACG
TATTTTCTAAGTTGGAAAGCAGGTTTTCTTAAAAAATTATAGACACGGTCACT
AAATTGATTAGTCAGAATTCTAGACTGAAAGAACCTAAACAAAAAAATTAAAGATA
TAAATATATGCTGTATGTGTTATGTAATTCTTAGGCTATAATACATTTCTTATTTCGC
ATTTCATAAAATGTCCTAATACAAAAAA

FIGURE 4

MVKFPALTHYWLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLAYSLMKFF
TGPMMSDFKNVGLVFVNNSKRDRTKAVLCMVVAGAIAAVFHTLIAYSDLGYYIINKLHHVDESV
GSKTRRAFLYLAAPPFMDAMAWTAGILLKHKSFLVGCAISIDVIAQVVFVAILLHSHLEC
REPLLIPILSLYM GALVRCTTLCGLYYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL
ATQRISRPIVNLFVSRLGGSSAATEAVAILTATYPVGHMPYGWLTEIRAVYPAFDKNNPSN
KLVSTSNTVTAAHIKKFTFVCMALSLTLCFVMFWTPNVSEKILIDIIGVDFAFaelCVVPLR
IFSFFPVPTVRAHLTGWLMTLKKTFLAPSSVLRIIVLVIASLVLPYLGVHGATLGVGSL
AGFVGESTMVAIAACYVYRKQKKKMENESATEGEDSAMTDMPPTEEVTDIVEMREENE

Transmembrane domains:

amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374,
408-423, 431-445

FIGURE 5

CCTGACAGAACGTGCCCGGAGCTGGGGAGATNCAACATTAAAGAAGATGCTGAGCTCTGGT
GCCNTTGGCTCTAATTCTGCCACACAGAGAACAGTCGGCTATTGTCAACCTTTGTT
TCCCGGGACCTTGGCAGTCTGCAGCACAGAGGCAGTGGCGATTTGACAGCACATA
CCCTGTGGGTACATGCCATACGGCTGGTGACGGAATCCGTGCTGTATCCTGCTTCG
ACAAGAATAACCCCAGCAACAAACTGGTGAGCACAGCAACACAGTCACGGCGGCCACATC
AAGAAGTTCACCTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTCGTGTGTTGGAC
ACCCAACGTGTCTGNAAAATCTTGATAGACATCATCGGAGTGGACTTGCCTTGAGAAC
TCTGTGTTGTTCCCTTGCGGATCTTCTCCTCTTCCAGTCCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTC

FIGURE 6

TGACGGAATCCGGGCTGGGTATCCTGGTTNGACAAGATAAACCCCCAGCAANAAATTGGG
GAGCAGGGAAAACAGTNACGGGCAGCCCACATCAAGAAGTTCACCTTNGTTGNATGGNTC
TGTCAACTCACGCTNTGTTCTGTGATGTTGGACACCCAAAGTGTTGAGAAAATTGAT
AGACATNATCGGAGTGGANTTTGCCTTGAGAANTTGNNTGTTCTTGCAGGATTTCT
CCTTTTCCCAGTTCAGTCACAGNGAGGGCGATCTCACCGGGNGNTGATGACANTGAAG
AAAACCTTGTCTTGCCTTGCCTTGCAGCTNTTGGTGCGGATCATTGTCCTNATNGCCAGCCTTGT
GGTCCTACCCCTACCTGGGGGTGCACGGTGCACCCCTGGCGTGGGTTCCCTCCTGGGGCA

FIGURE 7

TATTCCCAGTTCCGGTCACTGGGAGGGCGCATNTCACCGGGTGGCTGANGACACTGAAGAAA
ACCTTNGTCCTGCCCGCAGNTTGTGNTGCGGATNATCGTCCTCATGCCAGCCTNGTGTT
CCTACCCTACCTGGGGGTGCACGGTGAGAC

FIGURE 8

FIGURE 9

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLLIFTFFLFLGVLVSIIIMLSPGVESQLYKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSSRDPRAAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHWNQRWLGAEECDSRAYLAGLFFFLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFCVCVSIAAVLPKVQDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVA GPEGYETQWWADPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVA ACEGRAFDNEQDGVTYSYSFFHFCLVLASLHVMMTLMWYKPGETRKMISTWTAVWVKICSWAGLLLWTLVAPLLLNRNDRFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

09272544-101601

FIGURE 10

GAGCGAGGCCGGGACTGAAGGTGGGTGTCAGGCCCTGGCAGAGGGTTAACCTGGGTC
AAATGCACGGATTCTCACCTCGTACAGTTACGCTCCCGCGCAGCTCGCAGGGACTTGA
AGTCCTGAGCGCTCAAGTTGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGC
ACCGCGGAGCTTCTCTGTAGACGATTGCCCTATTCCCGAGCTTTGTGCCAGACTG
TGACTGCCGATTCGGAAGTCCTTGAGGAGCGTCAGAACGGCTCCCTACGTCCCAGGCC
TATTACCGGAATCTGGATGGGACCGCCTCCGGAGCTGTTGGCAAAGATGAACAGCAGAG
AATTCAAAGGACCTTGTCAATATCTGTAAAGACGGCAGCTACAGCAGGCATATTGGCTGGG
TGTATGGGAAATACCAGTTTATTCTGCTAACACAATACATTGAGCAGAGGCCAGGCA
GAAATTTATCATAACCGGTTTGATGCTGCAATCTGCACATCGTGTGCCACACGGAGCTT
CATCGTTGGCTGGCGCTGGGGTGAGAAACTGCGAGTGTGTGACTATATTCAACACAG
TGAACACTAGTCTGAATGTACCGAATAAAGATGCCCTAACGCCATTGGTAATTGAGGA
GCTGTCACGGAAAGTCTTTTAGATAAACGTTAGGCCGCTGGCTGGCTGGTGGCTGGCCAT
AATTGGAGCCTGCTGGCACTCCTGTAGGAGGCCGCTGATGGCATTTCAGAAGTACGCTG
GTGAGACTGTTAGGAAAAGAAAACAGAAGGATCGAAAGGCACTCCATGAGCTAAACTGGAA
GAGTGGAAAGGCAAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTACG
GGAAGGATGAACCTGAGAATGTGCTAAAGAAAATTGAAAGCACTGCTAAACCTCTAGAAACC
CTTCAGTAATAGATAAACAGACAAGGACTGAAGTGTCTGAACTGACTGGAGA
GCTGAAGGGAGCTGCCATGTCGATGAATGCCAACAGACAGGCCACTCTTGGTCAGCTGC
TGACAAATTAAAGTGTGGTACCTGTGGCTGGCAGTGGCTGCTTGTCTTTCTTCTTCTT
TTAACTAAGAATGGGGCTGTTGACTCTCACTTACTTATCCTAAATTAAATACACT
TATGTTGTATTAACTATCAATATGCAACATGGATATACCCACCTAGATTAA
GCAGTAAATAAAACATTGCAAAAGATTAAAGTTGAATTTCAGTTT

FIGURE 11

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318
><subunit 1 of 1, 285 aa, 1 stop
><MW: 32190, pI: 9.03, NX(S/T): 2
MEVPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRRLRELF
GKDEQQRISKDLANICKTAATAGIIGWVYGGIPAFIHKQQYIEQSQAEIYHNRFDAVQSAH
RAATRGFIRYGWRWGWRATAVFVTIFNTVNTSLNVYRNKDLSHVIAAGAVTGSLSRINVGLR
GLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEEWKGRQLQVTEHLPE
KIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD
```

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and 171-195

N-glycosylation site:

amino acids 153-156

FIGURE 12

CGGAAGTCCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCGGA
ATCTGGATGGGACCGCTCCGGGAGCTGTTGGCAAAGATGAACAGCAGAGAAATTCAAAGGA
CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGAA
TACCAGCTTTATTCACTGCTAAACAACATACTTGAGCAGGCCAGGCAGAAATTATCAT
AACCGGTTTGATGCTGTGCAATCTGCACATCGTGTGCCACACGAGGCTTCATTGTTCATG
GCTGGCGCCGAACC

FIGURE 13

TCAAGTTGTCGTAGTCGAGAGAAGGCCATGGAGGTGCCACCGCACCGCGAGCTT
TTTTCTGTAGAGCATTGTGCCTATTTCCCGAGTTTGCTGCCGAAGCTGTACTGCCAT
TCGGAAGTCCTTGAGGAGCGTCAGAACGGCTTCCCTACGTCCCAGAGCCCTATTACCGGA
ATTTGGATGGGACCGCCTCCGGAGCTGTTGGCAAAGATGAACAGCAGAGAATTCAAAGG
ACCTTGCTGATATNTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGA
ATACCAGCTTTATTGNTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTATNA
TAACC

FIGURE 14

GAGCCGCCGCCGCGCGCGCGCAGCTGCAGCCCCAGGCCCGCCCCACCCACGTCT
GCGTTGCTGCCCGCCTGGGCAGGCCCCAAGGCAAGGACAAAGCAGCTGTCAGGGAACCT
CCGCCGGAGTCGAATTACGTGCAGCTGCCGCACCCACAGGTTCAAGATGGTTGCGGGG
GCTTCGCGTGTCCAAGAACAGCTGCGCTGCCCTCAACCTGCTTTACACCTTGGTAGTCTG
CTGCTAATTGGAATTGCTGCGTGGGCATTGGCTCGGGCTGATTCCAGTCTCGAGTGTT
CGGCGTGGTCATTGCAGTGGCATCTTCTGTTCTGATTGCTTTAGTGGGTCTGATTGGAG
CTGTAAAACATCATCAGGTGTTGCTATTTTTATATGATTATTCTGTTACTGTATTATT
GTTCAAGTTTCTGTATCTGCGCTTGTAGCCCTGAACCAGGAGCAACAGGGTCAGCTCT
GGAGGTTGGTGGAACAAATACGGCAAGTGCCTGAAATGACATCCAGAGAAATCTAAACTGCT
GTGGGTTCCGAAGTGTAAACCCAAATGACACCTGCTGGTAGCTGTGTTAAAAGTGACCAC
TCCTGCTGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTGAGATTGTTGG
TGGCATTGGCTGTTCTCAGTTTACAGAGATCCTGGGTGTTGGCTGACCTACAGATA
GGAACCAAGAACCCCCCGCGAATCCTAGTGCATTCTTGTGAGAAAACAAGGAAGAT
TTCTTCTGTATTATGATCTGTTCACTTCTGTAATTCTGTTAGCTCCATTGCCAGT
TTAAGGAAGGAAACACTATCTGAAAAGTACCTTATTGATAGTGGATTATATATTACT
CTATGTTCTCACATGTTTTCTTCCGTTGCTGAAAATATTGAAACTTGTGGTCTC
TGAAGCTCGGGCACCTGAAATTACTGTATTCTTGGCTGGCACTGTCACGTGGCTT
TCTTAGCATTTCACCTGAGAAAACATTGTTAGGTAACACTGTGTGGTTATATGGTGAA
TCTGAACGTACATCTACTGGTATAATTATATGTAGCACTGTGCTGTAGATAGTCC
TGGAAAAAGAGTGGAAATTATTAAAATCAGAAAGTATGAGATCCTGTTATGTTAAGGGAAA
TCCAAATTCCAATTTTGGTCTTTAGGAAAGATTGTTGTGGTAAAAGTGTAGTA
AAAAAATGATAATTACTGTAGTCTTTATGATTACACCAATGTATTCTAGAAATAGTT
GTCTTAGGAAATTGTGGTTAATTGTGACTTTACAGGTAAGTGCACAGGAGAAGTGGTT
CATGAAATGTCTAATGTATAATAACATTACCTTCAGCCTCCATCAGAATGGAAAGAGTT
TGAGTAATCAGGAAGTATATCTATATGATCTTGTATTGTTTATAATAATTGAAAGTCTAA
AAGACTGCATTAAACAGTTAGTATTAAATGCGTTGGCCACGTACAAAAAGATATTG
ATTATCTAAAAATTGTTAAATACCGTTTGTGAAATTCTCAGTATTGTAACAGCAACTT
GTCAAACCTAACGATATTGAAATGATCTCCATAATTGAAATTGAAATCGTATTGTGTG
GCTCTGTATATTGTTAAAAAATTAAAGGACAGAAACCTTCTTGTGTATGCATGTTGA
ATTAAAAGAAAGTAATGGAAG

FIGURE 15

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979
><subunit 1 of 1, 204 aa, 1 stop
><MW: 22147, pI: 8.37, NX(S/T): 3
MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGGLISSLRVVGVVIAVGIFLFLIALV
GLIGAVKHHQVLLFFYMIILLLVPIVQFSVSCACLAQNQEQQGQLLEVGWNNNTASARNDIQR
NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRFVGGIGLFFSFTEILGVWL
TYRYRNQKDPRANPSAFL
```

Signal Peptide:

amino acids 1-34

Transmembrane domains:

amino acids 47-63, 72-95 and 162-182

FIGURE 16

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGTNATTTTTTATATGATTATTCTGTAANT
TGTATTTATTGTTCAAGTTNTGTATCTGCCTGTTAGCCNTGAACCGAGAACAGG
GTCAGNTNTGGAGGTTGGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAT
NTAAACTGCTGTGGGTCCGAAGTGTAAACCAAATGACACCTGTNTGGCTAGCTGTGTTAA
AAGTGACCACNTGTGCTGCCATGTGCTCCAATCATAGGAGAAATATGCTGGAGAGGTTTGA
GATTGTTGGTGGCATTGGCCTGTTAGTTACAGAGATCCTGGGTGTTGGCTGACC
TACAGATAACAGGAACCAAG

FIGURE 17

AATCCCCAATTCCCAATTGGCTTTAGGAAAGATGTGTGTGAAAGTGT
TAGTATAAAAATGATAATTACTGTAGTCCTTATGATTACACCAATGTATTCTAGAATAG
TTATGTCTTAGGAAATTGGTTAATTGGACTTTACAGGTAAAGTCAGGAAAGGAGAAGTG
GTTTCATGAATGTTAATGTATAAACATTACCTTCAGCCTCCATCAGAATGAAAGC
AGTTTGAGTAATCCAGGAAGTATATCTATATGATCTGATATTGTTTATATAATTGAAAG
TCTAAAAGACTGCATTAAACAAAGTTAGTATTAAATCGCTTGGCCACGTAGCAAAAGAT
ATTGATTATCTTAAATTGTTAAATACCGTTTCTAGAAAGTCTCAGTATTGTAACAGC
AACTTGTCAACCTAACGATATTGAAATATGATCTCCCATATTGAAATTGAAATCGATT
GTGTGGAGGAAATGCCATTATGTGTGCTGAAGGACACAGTAAGAGCACCAAGTTGTGCC
CCACTTGC

FIGURE 18

ATGATTATTCTGTTACTTGTATTTATTGTTCAAGTTTATGGTATCTTGCCTTGTAGCCC
CTGAAACCAGGAGCAACAGGGNNNCAGCTTCCTGGAGGTTGGTTGGCAACAATCACGCCAAG
TGACTCCGCAAATGACATCCCAGAGAAATCCTAAACTGCTGTGGGTTCCGAAGTGTAAACCC
AAATGACACCTGTCTGGCTNGCTGTGTTAAAGTGACCACTCGTGCTGCCATGTGCTCCAA
TCATAGGAGAATATGC

FIGURE 19

CAGTCACCATGAAGCTGGGCTGTCTCATGGCCTGGGCCCTCACCTTCCCTGGTG
CTCTGGGTGGCCAGATGCTACTGCGTCAGTTTGAGACGCTGCAGTGTGAGGGACCTGT
CTGCACTGAGGAGAGCAGCTGCCACACGGAGGATGACTGACTGATGCAAGGGAGCTGGT
TCCAGGTCAGGCCAACACTTCACTGGAACCTTCCACCTGATTGTGCTCTATGACTGGCTG
ATCCTCCAAGGTCAGGCCAACAGCTGGGCTTCACTGAGGAGCTGGTCAAAAGGGAGCTGGCTG
CTGGCAAGACTGGCCACTGACTCAAGTGACCTCTACCGAGATGGCTCAGCTCTGGTCCCC
CCGGGCTAACAGGGAAATTCTCATCACCGTGGTACAAAAGGGAGACAGCGGGCACTACAC
TGCAGTGCCATCTTCCAGAGGCCCTGGTCTGGGATCCAGAGAAACAGCATCTGGTGGCTAT
CACAGTGCCAAGAACGTTTCCAGGCCAACCTCAGAGCTGTACCCCTAGCTAACCCCAG
CAGGAAGCCCCATGACCTTGAGTTGTCAGACAAAGGTCGCCCTGCAGAGGTAGCTGCCGC
CTCCCTTCTCTTCTACAAGGATGGAAGGATAGTGCAAAGCAGGGGCTCTCTCAGAAATT
CCAGATCCCCACAGCTTCAGAGATCACTCCGGTCTACTGGTGTGAGGCAGGCCACTGAG
ACAACCCAATTGGAAACAGGCCAACAGTAGAGATCAGAGTGAGGTGCTTCCAGCTT
GCTGCACCTCCCACATTGAACTCAGCTCTCAGAAATCAGCTGCTCCAGGAACGCTCC
GGAGGGCCCTGGGCTCTGGCTCCGCCAACCTCATCTGGAGATGCCAGGCTTTCTT
CTCCCTGGGATGCCAGATCTCATCTGATACCCAGATGGGCTTCTCTCAAACACATG
CAGGATGTTGAGAGTCTCTCGGTACCTGCTCATGGAGTTGAGGGATTATCTGGCCACCA
GAAGCCTGGGACCAAAAGCTACTGTAAGAAGTAAACAGTTCATCCATGATCTCACT
TAACCACCCCAATACTGATTCTTATTTCTCTCTGCTCTGCAACATATGCAATAAGTA
CTTTTACAAGGTGCCCAGTTGTTGTTAGAAATGTTAGTTAGGTGAGTGAAATAAATT
ATATAAAAGTGAGAAATTAGAGTTAGCTATAATTGTTGATCTCTCTTAAACACAGAATT
TGCTGTCTAGATCAGGAATTCTCATCTGTTATATCGACCCAGAAATGTTGATTTAAAGAGAA
CTAATGGAAAGTGGATTGAATACAGCTCAACTGGGCCAATTGGCCCCCAGAGGACA
TTGGGCCATGTTGGAGACATTTCGTCATTATACTTGGGGGTTGGGGATGGTGGGATGT
GTGTCACGGCATCCAGTAATAAGGCCAGGGGTGCCCTAAACATCTATAATGCAACAG
GGCAGTACCCCCAACAGAAAAAAATCTGGCCCAAAATGTCAGTTGACTGAGTTGAGAAA
CCCCAGCCATATGAAACCTAGGTGTTGGGCTCTGGAAATGGGACTTGTCCCTCTAATT
TATCTTTCCAGCCTCATCAGTATTCTACTGACATACCAGTCTTAGCTGGTGTCT
GTCTGTTCTTAGTTGTTCTAGTTGTTGATCCCTCAAAAGGCCATTATGTTGAAATCTTAAATCCCC
AAGGTGATGGCATTAGAAGTGGGCTTGGGAGTGATTAGATCAGGAGTGAGAGGCCCTC
ATGATTAGGATTAGTGCCTTATTTAAAAGGCCAACAGAGCTAACCTCCACCAT
ATGAGGACGTGCCAGAAAGATGACATGTATGAGAACCAAAAACAGCTGTCGCCAACACCG
ACTCTGTCGTTGCCCTTGATCTTGAACTTCCAGCCAGAAACTATGAGAAATAAAATTCTGG
TTGTTGTAAGCCTAA

FIGURE 20

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594
><subunit 1 of 1, 359 aa, 1 stop
><MW: 38899, pI: 5.21, NX(S/T): 0
MKLGCVLMAWALYLSLGVLWVAQMILLAASFETLQCEGPVCTEESSCHTEDDLTDAREAGFQV
KAYTFSEPFHLLIVSYDWLILQGPAKPVFEGDLLVLRCQAWQDWPLTQVTFYRDGSALGPPGP
NREFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGS
PMTLSCQTKLPLQRSAARLLFSFYKDGRIVQSRLSSEFQIPTASEDHSGSYCEAATEDNQ
VWKQSPQLEIRVQGASSAAPTLNPAPQSAAPGTAAPEEAPGPLPPPPTPSSEDPGFSSPL
GMPDPHLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE
```

Signal sequence:

amino acids 1-17

Leucine zipper pattern sequence:

amino acids 12-33

Protein kinase C phosphorylation site:

amino acids 353-355

101501

FIGURE 21

FIGURE 22

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45416
><subunit 1 of 1, 321 aa, 1 stop
><MW: 35544, pI: 8.51, NX(S/T): 0
MGILLGLLLGHLTVDTYGRPILEVPESVTGPWKGDVNLPCTYDPLQGYTQVLVKWLVQRGS
DPVTIFLRDSSGDHIQQAKYQGRLHVSHKVPGDVSLQLSTLEMDDRSHYTCETWQTPDGNQ
VVRDKITELRVQKLSVSKPTVTTGSGYGFVPGMIRISLQCQARGSPPISYIWYKQQTNNNQE
PIKVATLSTLLFKPAVIADSGSYFCTAKGQVGSEQHSDIVKFVVKDSSKLLKTKEAPTTMT
YPLKATSTVKQSWDWTTDMGDYLGETSAGPGKSLPVFAIIILISLCCMVVFATMAYIMLCRKT
SQQEHVYEAAK
```

Signal Sequence:

amino acids 1-19

Glycosaminoglycan attachment site:

amino acids 149-152

Transmembrane domain:

amino acids 282-300

FIGURE 23

GGCGCCGGGAGCCCCATCTGCCCCAGGGCAGGGGCGGGGGCGGGGCTCCGCCGCACAT
GGCTGCAGCACCTCGCGCACCCCGAGCGCCGCCAGCTGCCAGGTCCCTCGGA
GGCGCCGGCGGCCGCCAGCCAAGCAGCAACTGAGCGGGAGCGCCGTCGGGAGTC
GGGATGTCCTCTCCCTCTCCCTTCTGCTAGTTTCTCATATGTTGAACTTGGGAGCTCA
CACTGAGATCAAGAGAGTGGCAGAGAAAAGGTCACTTGCCACCATCAACTGGGC
TTCCAGAAAAAGACACTCTGGATATTGAATGGCTCTCACCGATAATGAAGGAAACAAAAA
GTGGTGATCACTTACTCCAGTCGTCACTGCTACAATAACTTGAACAGAGGGCCG
AGTGGCCCTTGCCTTCAATTCTGGCAGGAGATGCCCTTGCGAGATTGAACCTCTGAAGC
CCAGTGATGAGGGCGGTACACCTGTAAAGGTTAAGAATTCAAGGGCCTACGTGTGGAGGCCAT
GTCATCTAAAGTCTTAGTGGACCATCCAAGCCAGTGTGAGTTGAAAGGAGAGCTGAC
AGAAGGAAGTGAACCTGACTTTGCACTGTGAGTCATCTCTGACAGAGCCCATTGTTGATT
ACTGGCAGGAATCCAGAGAAAAGGGAGAGGAGGAGGAAACGTCCTGCCCTCCAAAATCTAGGATT
GACTAACACCCTGGACGAGTTCTGCTGCGAGATACTTACCATGTCCTACTCTGAGCTGA
CCAGTGCACAGCAGGCAAAGAAGCTGGAGGAAAGCTGTGTGCGAGTAACCTACAGT
ATGTCAGGAGCATCGGCATGGTGTGAGGAGCAGTGACAGGCTAGTGGCTGGAGGCCCTGTG
ATTTTCTCTTGGTGTGCTAATCGAAGGAAAGAACAAAGGATATGAGGAAGAAGA
GAGACCTAATGAAATTGAGAAGATGCTGAAGCTCAAAAGCCGCTTGTGAAACCCAGCT
CCTCTTCTCAGGCTCTGGAGCTCACCGCTCTGGTTCTCCCTCCACTCGCTCCACAGCAAAT
AGTGGCTCACGCAGGGCACAGTCAACTGACGCGAGCAGCCAGGGCTGGCCAC
CCAGGCATACAGCCTAGTGGGGCAGAGGTGAGAGGTCTGAACCAAAGAAAGTCAACCATG
CTAATCTGACCAAGCAGAAACACCCAGCATGATCCCCAGCAGAGCAGAGGCCCTTCAA
ACGGCTGAATTACAATGGACTTCACTGCCAGCTTCTAGGAGTCAGGGCTTTGACTC
TTCTCGTCATTGGAGCTCAAGTCACAGGCCACAAACACAGATGAGGGTCACTAAGTAGCA
GTGAGCATTGACGGAACAGATTAGATGAGCATTTCTTATACAAATACCAAAAGCAA
AGGATGTAAGCTGATTCACTGTAAAAGGCATCTTATGTCCTTGTGACAGAGTAAGGG
AAAGCAGGAGTCCAAATCTATTGTTGACAGGACCTGGTGTGAGAAGGTTGGGGAAAGGTG
AGGTGAATATACTTAAACTTAAATGTTGGATATTGTATCTAGTGTCTTGTGATTCACAATT
TTCAAGAGGAATGGGATGCTGTTGIAAATTCTATGCAATTCTGCAACATTATTGGATT
ATTAGTTATTGCTGACAGTCAGCAGAACCCACAGCCTTATTACACCTGCTACACCATGTC
TGAGCTAACACTCTAAGAAAACAGGAAATTCTGAGGAAAGGGATCTAGGAATGCTGAAGATTA
TTCATTGTCATAAGGTTGGATATTAAATTCAAGGGAGTTGAAATAGTGGGAGATGGAGA
AGAGTGAATGAGTTCTCCACTCTATACTAATCTCACTATTGTATTGAGGCCAAATAAC
TATGAAAGGAGACAAAATTGTCAGAACAAAGGATTGAGAGCTTCCATCTTGTGATGTT
ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTTGATTCCCTCAAAT
CAGATGCCCTCAAGGACTTTCTGCTAGATATTCTGAGGAAAGGAAATACACATGTCATT
TATCAACGTCCTTAGAAAAGAATTCTTCTAGAGAAAAGGGATCTAGGAATGCTGAAGATTA
CCCAACATACCAATTATAGTCCTTCTTCTGAGAAAATGTGAAACACAGAAATTGCAAGGACTGG
GTGGAAGTGAAGGGAGATTAGATCAGTTCTCTTAAATATGTCAGGAAGGTAGCCGGCA
TGGTGCCAGGCACCTGTAAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTC

FIGURE 24

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419
><subunit 1 of 1, 373 aa, 1 stop
><MW: 41281, pI: 8.33, NX(S/T): 3
MSLLLLLLVSYYVGTLGHTTEIKRVAEKVTLPCHHQLGLPEKDTLDIEWL LTDNEGNQKV
VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPCELEGELETGSQTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID
YNHPGRVVLLQNLTMSSYSGLYQCTAGNEAGKESCVVRTVQYVQSIGMAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEERPNEIREDAEAPKARLVKPSSSSSGSRSRSSRGSSSTRSTANS
ASRSQRTLSTDAAQPQGLATQAYSLVGPEVRGSEPKVHHANLTKAETTPSMIPSQSRAFQTV
```

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

135978544.101503

FIGURE 25

FIGURE 26

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52594
><subunit 1 of 1, 655 aa, 1 stop
><MW: 71845, pI: 8.22, NX(S/T): 8
MGTSPSSSTALASCRIARRATAMIAGSLLLLGFLSTTAQPEQKASNLIGTYRHVRDRTG
QLTCDKCPAGTYVSEHCTNTSLRVCSSCPVGTFTRHENGIKECHDCSQPCPWPMIEKLPCA
ALTDRECTCPPGMFQSNTCAPIHTVCPGVGVRKKGTETEDVRCKQCARGTFSDVPSSVMKC
KAYTDCLSQLVVIKPGTKETDNVCGTLPSFSSTSPSPGTAIFPRPEHMETHEVPSSTYVP
KGMMNSTESNSSASVRPKVLSSIQEGLVDPNTSSARGKEDVNKTLPNLQVNNHQQGPBHRHIL
KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFIDINEHLPMIVLFLLLVLVVIVVCSIRK
SSRTLKKGPQRQDPSAIVEKAGLKKSMTPQNREKWIIYCYCNGHGIDILKLVAQVGQSQWKDIY
QFLCNASEREVAAFSNGYTAGHERAYAALQHWHTIRGPEASLAQLISALRQHRRNDVVEKIRG
LMEDTTQLETDKLALPMSPSPSPSPPIPSPNAKLENSALLTVEPSPQDKNKGFVDESEPLL
RCDSTSSGSSALSRNNGSFITKEKKDTVLRQVRLDPCLDQPIFDDMLHFLNPEELRVIEEIPQ
AEDKLDRLFEIIIVGKVSQEASQTLLDSVYSHLPDLL
```

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 350-370

FIGURE 27

ATGGGAGCAGTAACACTGTGGCTACTATCTTCTCGTGGGCCATCTACATTTGGGA
CTCGGAAATTATGAGCTAGGGCTGGAGCCGAGTCAGGGTCTGAAATAGTCAC
CATGGGGAAAATGATCCGCTGTGTAAGGCCCCCTCTCATTCGATCCTTTGGCC
TTGTATGATTGAAAATAAGTCTGTGCAACAGATGCAGATGCTGTGCAAGATCCTG
TCACGTGCCATTGAGTTTCTAACATCGTCATTGGATCATTGATTGATATTAGC
ACTGGCCATTGCTCGGCATCACCTGCACTGCTCAGGGAACTACAGATGCTGCTCATCT
TTAAGTGATCTGAGCTGATAGCTGATGTCAGGACTCTCGGATGCAAGAGCGGGAGGC
GAGTACCGCTGTGTCGGGGTGGTGGTCAGAACTCCGGTGTCCAGGTGTTCACAGTGCTTC
GTGGAAGACCATGTCCTGGATGACTGAAAGGGTCACTACGGAATGTTGCTGTGCCAAC
TGGGTTTCCAAGCTATGTGAGTTCAGATAACCTGAGTGCAGCTGCTGGAGGGCAGTTC
CGGGAGGAGTTGTGTCATCGACCATCTTGCAGATGACAAGGTGACTGCATTACACCA
CTCAGTATATGTGAGGGAGGGATGTGCCCTGGCACCTGGTTACCTTGCACTGCACAGCCT
GTGGTCATAGAAGGGCTACAGTCACGCCATCTGGTGGGAAATACGTCTCTGCTCGAG
TGGCCCTGGCAGGCCAGCCTTCAGGTCACGGGCTACCCAGTGTGCGGGGCTCTGTAC
GCCCTCTGGATCATCACTGTCACACTGTGTTATGACTGTACCTCCCCAAGTCATGGA
CCATCAGGGGGCTAGTTTCTGTGAACTACAGGCCCCATCCCCTATGGTGAGAAG
ATTGTCTCACACAGCAAGTACAAGGCCAAAGGGCTGGCAATGACATGCCCTTATGAGCT
GGCCGGGCCACTCTTCATAGAATGTCAGCTGCTGCTGGCAACTCTGAGAGA
ACTTCCCGATGAAAAGTGTGCTGGACGT CAGGATGGGGGCCACAGGGATGGAGGTGAC
GCCTCCCTGCTCTGGTGAACACCAGGGGGCTCCCTTGTGTTTCAACAAAGTCGACAC
GGACGTGTAAGGTGGCATCATCTCCCCCTGCTGTGCGGCCATCTGACGGTGGC
TGGACAGCTGCCAGGGGAACAGGGGGCCCTGGITGTCAAGAGAGGGCTGTGGAAG
TTAGTGGGAGCGACCAAGCTTGGCATCGCTGCGAGGGTAAACAGCTGGGGTGTACAC
CCGGTCACTCTCTCTGGACTGGTACCGAGCAGATGGAGAGACCTAAACCTGG
GAGGAAGGGGACAATGAGCCACTGGTCTGGAGGTGTGAAAGACAGCCGATCTCT
GGACTCCCGTGTAGGAACCTGACACAGCAGACACCTTGGAGCTGTAGTCCGGCACC
GTAGCAGGGCCAGAGGACCACTTCCATCTGATTCAGGACACAACTTCAAGCTGT
GTTTTTGTGTTTGTGAGGTGGAGTCTCGCTGTGCCCCAGGCTGGAGTCTGGGGAAA
TCTCTGCTACTGAGCCTCCGCTCCGTTCAAGCGATTCTCTGCTCAGCTTCCCCA
GTAGCTGGGACCAAGGTGCCCCCAGACACACCCAACTAATTGTTGTTAGTAGAGAC
AGGGTTTCCACATGTTGGGAGGCTGCTCAACCCCTGACCTCAATAGTGTGCTGCTT
CAGCCCTCCACAGTGTGGGATTACAGGGTCAAGGCCACAGCCTAGGCTCAGCTCTT
TGATCTTCAAGAACAAAAGCAGCAACTTGCAGGGGGCTTCCACTGGTCCAT
CTGGTTTCTCTCCAGGGTCTGGCAAATTCTGACGAGATAAGCGATTGTGACCTCAG
CAAGGCCACCAACAGCCACTGCAAGGGCAGCACCAGGCCAAAGTGCAGAACCTGAGC
ACTGCACCTTTCTACCTCTAGGGACCAAGGACCAACAAACCCCTTCTACTTCAAGACTTAT
TTTACATGTGGGGAGGTTAATCTAGGAATGACTCTTAAAGGCTTATTTCATGATTCTT
TGTGAGTCTGGGCTGTGAGCTATTGTCTTGTGATCCAAATAATGTTCTTCCCT
CATGGTCTGGGCTGTGAGCTGTGAGTCAATGAAATCATCCACTGAAA

FIGURE 28

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234
><subunit 1 of 1, 453 aa, 1 stop
><MW: 49334, pI: 6.32, NX(S/T): 1
MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILA
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVRVGGQNAVLVQVFTAAS
WKTMCSDDWKGHYANVACAQLGFP SYVSSDNLRVSSLEQQFREEFVSI DHLLPDDKVTAHH
SVVREGCASGHVVTLQCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVIT
PLWIITAHCVYDLYLPKSWTIQVGLVSLLDNPAPSHILVEKIVYHSKYKPKR LGNDIALMKL
AGPLTFNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAPLISNKICCNHR
DVYGGIISPSMLCAGYL TGGVDSCQGDGGPLVCQERRLWKLVGATSGFGIGCAEVNKG VYT
RVTSFLDWHHQMERDLKT
```

Signal Peptide:

amino acids 1-20

Transmembrane domain:

amino acids 240-284

FIGURE 29

CCACAGCGTCCGCTCTAGCCCCGGGCCAACTCGGACAGTTGCTCATTTATTCAGCAACGGCTGGCTGT
CCCAGAACGGCGCGCGCGCAGCACACACAGGGGAAACTTTTAAATAAAGGCTAGAAGA
GCTCAGCGCGCCGGCGCGCGAGGGCTCGCGAGGTGACTCCGGCAGGAGGAATCCCTCCGGCTCGCGA
CGCCCGGCCCGCGCTCGGGCGCCGCTGGAGGTGACTCGAGCGTCCGGCGCGGCCAGAGCTGCTGCACTGAAG
GCCGGCGACGATGCGAGCGCCGCTGCCCTGTCGCCCGCCGCCCTCTGCTCGCCCTGCGCGGTGCTCT
GCTCGCGGCCCTCGAGGGGCCGAGGGTGAGCTTATGAAAGGAGCTGATGAGTTGCTAGTGCCTCTGT
TCGGAGTGGGACCTCTGGATCCAGTGAAGAGCTCGACTCCAAGAATCATCCAGAGTGTGTAATATTGACT
ACAACGGAAAAGCAAAAGAACATGATCATAAATCTGGAAGAAATGAAAGTCTCATTTGCCAACGAGTTTACCGGAAAC
CCAACTATCTGCAAGACGGTACTGATGTCGCTCGAGCTGTTCTGGTCTCAGGGACTTATGTTGAAATATGTA
ACGGGGATATTCTGATTCAGCAGTCAGTCAAGCACGTTCTGGTCTCAGGGACTTATGTTGAAATATGTA
AAGCTATGTCCTAGAACCAATGAAAGTCAACACAAACGATACAACTTCCAGGGAAGAGCTGAAAGAC
CGGGGATCATGGATCACATCACACACCAACCTCGTCAAGAATCTGTTTACACCCCTCTCAGAC
ATGGGCAAGAAGGATCAAAGAGACCTCAAGGAACTAAGTATGTTGAGCTGGTGTGAGCTGCAAGACACCG
AGAGTTTACAGGGAAAGCTGGGAAAGATCTGGGAAAGATCTGGGAAAGGCTGGTGTGAGCTGCAATTCAGCTGACAA
GTTTACAGACCACTGAACATTGGATCTGTTGTTGGAGGAGTGTGAGATGACATGGAAATGCTGTCG
AAGCTGAGGACCTCAGGGCTTCAAGGAGTGTGAGCTGGGAAAGGAGGAGCTGGTGTGAGCTTCTACCTCGCAATCCA
TGACATCGCGACTTGTGAGCTGGGTTTATTCAGGGGACCATCGCATGGCCCAATCATGAGCATGTG
CACGGCAGACGGACTGTGGGGAAATGTCATGGGACATTGAGACATACTCCCGGACCTGGTGTGAGCTGGG
TGAGCTGGGCCCAATTCGGGATGAAATCATGACACACTGGAGCAGGGCTGTAGCTCTCAATGGGTTGAGAA
AGGAGGCTGCATCATGAGCGCTTCCACGGGTTACCCATTTCCTGGTGTGTCAGCAGTTGAGCAGGAGGAGCTT
GGAGACAGCGCTGGAGAACGGAGTGGGGTGTGCTGTGTTAACCTGCGGAAGTCAAGGGAGCTTCTGGGGCCA
GAAGTGTGGAGAACGATTGTGGAAAGAAGGAGGAGGTGTGACTGTGGGGAGGAGGAATGTATGATCGCTG
CTGCAATGCCGACCCACTGGTACCTGGGAGCTGGGAGCTGGGACATGGCTGTGAGACTGCGCACTGG
GAAGCTGAGGAAACAGCGTCAAGGACTCTGGGACATCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGG
TCACCTGGCCACCCAAATGTGTACCTGGAGCTGGGACACTGTGAGCTGGGAGCTGGGAGCTGGGAGCTGG
CTGGGAGACTCTGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGG
GAGAGTCAATTCTGGAGTGAATCTGGCAACTGTGGCAAAGTCTGAGAGTCTTGGCAATGAGAT
GAGAGCTTAAATGGGAAATCTGGAGCTGGGAGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGG
CATGAGAAACAAACATCCCTCTGAGCAAGGAGGAGGGGATTCTGGGGGGGACCAAGCTGACTCTGGGGATG
CATGGCAGGACCCAGGGCTTGTGCTGGAGCCAAACTGTGAGCTGGGAAATCTGGCTGAATGTCATG
AAATAATTAGTGTCTGGGGGTTACAGAGTGTGCACTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGG
CTGGGAGACTCTGGGAGCCACTGGGACCTCTCTGAGCAAGTGTGGCTTTGGAGGAAGGACAGACAGCG
CATGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ATCGCGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ATCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
CTGGCTCGAGAGGAGGTCAAGGGTCCCGAAAGGCCCTCTGTGACTCGAGCATTTGACTCTGGCTTTGGCATGTT
TCCATGAGAACACACACAGCTCTGGGCTCAGGGAGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGG
CAAGTGCAGGAGGGAGGGAGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGG
AGAGTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ACTGAGCTTCCACACAGCTGGGGGGAGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
TGCGAGCCCTGATGACTGGTCTGGCTGCAACTTAATGCTCTGATATGGCTTTAGCATTATTATGAAAT
AGCAGGGTTTATGTTAAATTATGAGACAGCTGGCCACCACTTCCATCTCCATCAGAAACTGAAATGGCAA
TGAAACAAACTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
AGTACTCAGGTTTGAGGGTTTGAGAGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
TGTTAAAGTGAAGACCAATGTAAGGCGCTAACTCCATCCCCGTGGCATTACTGCAAAAATAGAGTGCATT
GAAT

FIGURE 30

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624
><subunit 1 of 1, 735 aa, 1 stop
><MW: 80177, pI: 7.08, NX(S/T): 5
MAARPLPVSPARALLLAGALLAPCEARGVSLWNQGRADEVVSVASVRSGLWIPVKSFD SK
NHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTGH CYYHGV RG
YSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNR YKLFPAKKLKSVRGSCGSHH NTPNLA
AKNVFPPPSQTWARRHKRET LKATKYVELVIVADNREFQRQGKD LEVKQRLIEIANHVDF K
YRPLNIRIVLVGVVEVWNMDKCSVSDPFTSLHEFLDWRKM KLLPRKSHDNAQLVSGVYFQ G
TTIGMAPIMSMCTADQSGGIVMDHSNDNPLGA VTLAHELGHNFGMNHD TLDRCSCQMAVEK
GGCIMNASTGYPFPMVFSSCSRK DLETSLEKG MVGVCLFNLPEVRES FGGQKCGNRFVEEGEE
CDCGEPEECMNRCCNATTCTLKPDAVCAHGLCCEDCQ LKPAGTACRDSSNSCDLP EFCTGAS
PHCPANVYLHDGHSC QDV DG CYNGICQTH EQQC VTLWPGAKPAPGICCFERVNSAGDPYGN
CGKVSKSSFAKCEMRDAKCGKIQCQGGASRPVIGTN AVSIE TNIP LQQGGRILCRGTHVYL G
DDMPDPGLVLAGTKCADGKICLNRQCQNI SVFGVHECAMQCHGRGV CNNRKNC HCEAHWAPP
PCDKFGFGSTDSGPIRQAEARQEEAESNRERGQQQE PVGSQEHASTASLT LI
```

Signal peptide:

amino acids 1-28

FIGURE 31

TCCCCAAGGCTTCTGGATGGCAGATGATTNTGGGGTTTGCAATTGTTCCCTGACAACGAAA
ACAAAAACAGTTTGGGGTTCAGGAGGGAAANTCCAGCCTACCCAGGAAGTTGCAGAAACA
GTGCAAGGAAGGGCAGGANTTCTGGTTGAGNTTTGNTAAAACATGGACATGNTTCAGTG
CTGCTCNTGAGAGAGTAGCAGGTTACCACTTTGGCAGGCCAGCCCTGCAGCAAGGAGGA
AGAGGACTCAAAGTTGGCTTTCAGTGACACCTCAGCCTGGCAGCCCTGATAACTGGTNTNT
GGCTGCAANTTAATGCTNTGATATGGCTTTAGCATTATTATATGAAAATAGCAGGGTTTT
AGTTTTTAATTATCAGAGACCCTGCCACCCATTCCATNTCCATCCAAG

FIGURE 32

CATCCCTGCAACATGGTGAACACCACGCCCTGGCTAATTGGTTGTTGCTTGGTAGAGATGGGA
TTTCACCGTGTTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCGCCTCGGCCTCCC
AAAGTGTGGATTACAGGCAGTGCAACACCACCCGCCACAAACTTTAAGAAGTAAAT
GAAACCATACTTTACATTTAATGACAGGGAAAATGCTCACAAATAATTGTTAACCAAA
TTCTGGATAACAAAGTACAATCTTACTGTGTAACATGTATATGTACTATATGAAAATA
TACCAAATATCAATAACTTATCTCTGGGTAACCCCTCTCATACCCGTGCTAACAA
CTTTAACAAAAAATTGATCACTTTAACAGAAAATTCTGAAGGTATATGGG
ACAGAAAAAAACCAAGGGAAAATCACGCCACTTGGGAAAAAAGATTGAAATCTGCCT
TTTTATAGATTGTAATTAAAGGTCCAGGCTTCTAAGCAACTTAAATGTTGTTGCA
AACAAAGTACTGTCTGGATGTAGGGAGGAAAGGGAGTGATGTCAGTCCATTATGATGCC
TTGAATATAAGACCCACTTGCTATCTCCCTGCACAGCAGGAGCCACCCATCTCCAGC
ACACTGAGCAGCAAGCTGGACACACGGACACTGATCCAATGGTAAGGGATGGTGGCGA
TGCTCATTCTGGGTCTGCTACTTCTGGCGCTGCTCCTACCGTGCAGGTTTCTTCATTGTT
CCTTTAACAGTATGCCGGAAGCTACTGCAGCCAAACCAAAGCCCTCCAACAGTGCCT
ACAGCCTACAGCCGGTCTCTTGTGGCTTGCTTGCCTTCTACATCTTACAGG
CAGGTCAAGAACAGCTACAGTTCTCAACCCATACACTAAACCGAATCCAATGGTGCCT
AGAAGTTCAATGTGGCAAGGAAAAAAACAGGTCTTCAAAATCTACTAATTCTCCT
ATTAAACAGAGAACGCTTGAGAGTCTCAACTGGACTGGTTAAAGAGCATCTGAAGGATT
GACTAGATGATAATGCCGTACTCCCAGTACTTGGAGGCCCTAGGCCGGGATCACCTG
AGGTCAAGGAGTTGAGACTAACCTGGCCAAATGGTAAACCCATCTGACTAAAAATACA
AATATTGACTGGCGTGGTGGTGAAGTGCCTGTGATCCAGCTACTCAGGTGGCTGAAGCAGG
ACAATCACTGAACCTCAGGAGGCAGGGTTGCAGTGAGCTGAGATCGCCTACTGCACTCTA
GCCTAGCCTGGCAACAGAGTGAAGACTTCTGCTCAAAAAAAAAAGCCAAGTGCAGGGT
CACGCCCTGTAATCCCGGCACCTTGGAGGCCAGGTGGCGGATCACGGGTCAAGGAGATCA
AGACCATCTGGCTAATACAGTGAAACCCCTGTCTACTAAAAATACAAAAAATTAGCCGGG
GATGGTGGCAGGCACCTGGAGTCCAGCTACTCGGGAGGCTGAGGCAGGAGAATAGCGTGAA
CTCAGGAGGCCAGGAGCTTGCAAGTGAGCCGAGATTGCGCTACTGCACTCCAGCCTGGCGACAG
CGCGAGACTCCGTCTCAAAAAAAAAAAAAAA

FIGURE 33

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48309
><subunit 1 of 1, 67 aa, 1 stop
><MW: 6981, pI: 7.47, NX(S/T): 0
MGKGMVAMLILGLLLLALLLPVQVSSFVPLTSMPEATAAETTKPSNSALQPTAGLLVVLLAL
LHLYH
```

Signal peptide:

amino acids 15-27

FIGURE 34

FIGURE 35

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776
><subunit 1 of 1, 432 aa, 1 stop
><MW: 47629, pI: 5.90, NX(S/T): 0
MPARPGRLPLLLPAAALTALLLGHGGGGRGARQEAAAAADGPPAADGEDQDPHS
KHLYTADMFTHGIQSAAHFVMMFFAPWCGHCQLQPTWNDLGDKYNSMEDAKVYVAKVDCTAH
SDVCSAQGVGRGYPTLKLFKPGQEAVKYQGPRDFQTLENWMLQTLNNEEPVTPEPEVEPPSAP
LKQGLYELSAASFELHVAGQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCT
QHYELCSGNQVRGYPTLLWFRDGKKVDQYKGKRDLESLEYVESQLQRTEGTATETVTPSEA
PVLAEEPEADKGTVLALTENNFDTTAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA
GVKIAEVDTCAERNICSKYSVRGYPTLLLFRGGKKVSEHSGGRDLDLHRFVLSQAKDEL
```

Signal sequence:

amino acids 1-32

FIGURE 36

CTTTTCTGAGGAACCACAGCAATGAATGGCTTGATCCTGCTTCGAAGAAACCAATTAT
CCTCCTGGTACTATTCCTTGC~~AA~~ATTCAAGACTCTGGGTCTGGATATTGATAGCCGCTTA
CCGCTGAAGTCTGTGCCACACACAATTCACCAGGACCCAAAGGAGATGATGGTAAAAA
GGAGATCCAGGAGAACAGGGAAAGCATGGCAAAGTGGGACGCATGGGCCGAAGGAATTAA
AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCATTGGGAAGA
AGGGTGACAAAGGGAAAAGGTTGCTTGAATACCTGGAGAAAAGGCAAAGCAGGTACT
GTCTGTGATTGTGGAAGATACCGGAATTGTTGGACACTGGATATTAGTATTGCTCGGCT
CAAGACATCTATGAAGTTGTCAGAATGTGATAGCAGGGATTAGGAACTGAAGAGAAAT
TCTACTACATCGTGCAGGAAGAGAACTACAGGGAACTCCCTAACCCACTGCAGGATT~~CGG~~
GGTGGAAGTCATGCCATGCCAAGGATGAAGCTGCCAACACACTCATCGTACTTGTTGC
CAAGAGTGGCTTCTTCCGGTGTCATTGGCGTGAATGACCTTGAAAGGGAGGGACAGTACA
TGTCCACAGACAACACTCCACTGCAGAACTATAGCACTGGAAGTGAGGGGAACCCAGCGAC
CCCTATGGTCAGTGAGGACTGTGTGGAGATGCTGAGCTGCTGGCAGATGGAATGACACAGGT
CCATCTTACCATGTATTGCTGTGAGTTCATCAAGAAGAAAAGTAACTCCCTCATCCT
ACGTATTGCTATTTCCTGTGACCGTCATTACAGTTGTTATCCATCTTTTTCTG
ATTGTACTACATTGATCTGAGTCAACATAGCTAGAAATGCTAACTGAGGTATGGAGCCT
CCATCATCAAAAAAAAAAAAAAA

100978344 * 101601

FIGURE 37

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980
><subunit 1 of 1, 277 aa, 1 stop
><MW: 30645, pI: 7.47, NX(S/T): 2
MNGFASLLRRNQFILLVLFLIQSLGLDIDSRTAEVCAHTTISPGPKGDDGEKGDGPGEEG
KHGVGRMGPKGIKGELGMDQGNIGKTGPIGKKGDKGEEKGKAGTVCDCGRY
RKFVGQLDISIARLKTSMKFVKNVIAGIRETEEKFYIVQEEKNYRESLTHCRIRGGMLAMP
KDEAANTLIADYVAKSGFFRVFIGVNDLEREGQYMSTDNTPLQNYSNWNEGEPSDPYGHEDC
VEMLSSGRWNDTECHLTMYFVCEFICKKKK
```

Signal peptide:

amino acids 1-25

099785344 101501

FIGURE 38

GGTTCTATCGAATTGCCAACACTGGCCGATCTCTAGAGATCCTTCGACCTCGAC
CCACCGCGTCGCTGCTCTCCGGCGTGTGGAGTGGTGGGGGCTGGTGGAATGGCGCTGT
CGGCAGCGCACGCCGCTCTGGAGAGGAAGTCAGCTAGAAGCAGGCCCTAGGGTTT
CGGAAGGGAGGATCAGGGATGTTGCGAGCGGGCTGGAACCAGACGGTCCGATAGAGGAAGC
GGGCTCCATGGCTCCCTCTGCTGTCGCCCCCTGCTGCTGTGCTACCTGGCTGCTGCTGTA
AGCTAACCTCTGGCGCATGGCTGGCTCTGGCTTCCGGGACTTGGCTTGGCTGGAGCT
CTGTCGTCGAAAGGCTCTTCAGCTCGGCCCTGGCGCTGGCGCAGCAGCGCAGCGC
ACACCTTCTCATTCACGGCTCGGCCGCTTGGACTACAGAGGGAGCGCAGAGTAAC
AGGGCTCGACGCCCTCTACGTGCGTAGGGACTGGGACCCGAGCGGCCGACAG
CGGGCAGGGGAGCGCTGGAGAAGGGAGCGGGCAGCGCCGGAGCGGAGATGCA
GAAGCGGGCGGGAGTTGGCGAGGGAGGGTGCAGCGCAGAGGTGGAGGAGCGCCGGCC
CTGTCACCTGGCAACTGTGGCGCTGCTCCTCCCGCTGGCCAGAGTCTGTGGCTCTG
FTTCGGCTGGCAAGGCCGCCGCTGGCGACTGCTCTTGTGCCAACCGGCCCTGGCG
CCCTGCTGACTGCCTCCGAGCTGCGGCCGCGCGCTGGTGTGGCCAGAGTTCTG
GAGTCTCTGGAGCGGCCAGCTGGCCGGCTGGAGACGGCTGGGCTTCACTGGTGG
CCAGGAACCCCACCTCTGCTGGAATTAGGGATTCTGTGGCTGAGATGCTGGCTGAG
GCCAGTCCAGGATACCTCTCTCCCCAGAGCATAACAGACAGCTGCTGTACATCTT
ACCTCTGGCACCCAGGGCTCCCCAAGGTGCTCGATCAGTCATCTGAAGATCTGCAAT
CCAGGGCTTCTATCAGGTGTGGTGTCCACAGGAAGATGTATCACCTGCCCTCCAC
TCTACCATAGTGGCTTCTCTGGGATCCTGGCTGATGGCCATGGGCCACTGGGCA
GTGCTGAATCCAAGTTCTGGCTGGTCACTCTGGGAAGATTGCCAGCAGCACAGGTGAC
GGTGTCCAGCATATTGGAGGACTGTGCTGGATACCTCTGTCACCCAGCCCCGGAGCAAGG
AACGTGGCCATAAGTGGCCCTGGCAGTGGCGGCCAGGGCTGCGCCAGATACTGGAGCGT
TTTGTGCGGCCTTCGGGCCCCCTGCAAGGTGCTGGAGACATACTGGACTGACAGAGGG
GCCACCATCAACTACACAGGACAGCGGGCGCTGTGGGGCTGCTCTGGCTTACAAGC
ATATCTCCCTCTCTCTGATGTCAGTGCACACAGAGGCCAAATTCTGGGAC
CAGGGGACTGTATGCCACATCTCCAGGTGAGGCCAGGGCTGTGGTGGGGCGTGA
GCAGTCCCCATTCTGGCTATGCTGGGGCCAGAGCTGGCCAGGGGAAGTTGCTAAAGG
ATGTCCTCCGGCTGGGAGTGTCTTCTCAACACTGGGGACTGTGCTGTGCGATGACCAA
GGTTTCTCCGGCTTCTCATGATCTGACTGGAGACACACTTCTCAGGTGAGAGGG
CACAACCGAGGTGGCAGAGGTCTCGAGGCCCTAGATTCTCAGGAGGTGAACGTCTATG
GAGTCACTGTGCCAGGGCATAGAGGCCAGGGCTGAATGGCAGCCCTAGTTCTGCT
CACGGCTTGGGACCTTATGCACTCACACAGGTCTGAGAACTTGCACCTTATGCCG
GCCCGGATCTTCAGGCTCAGGAGTCTTGGCACACAGAGACCTTCAACAGCAGAAAAG
TTCGGATGCCAAATGAGGGCTCTGACCCAGCACCTGTCGACCCACTGACGTTCTGG
CAGGGCTGAGTGGCTACTCTGGGCCACTCAACACTGCCGGTACAGGCCCTCTGGCAG
CTTCGAACTTGAGAACCTTACACCTGGCAGCTGGAGAGGAAACTCTGTGGGGTGGGG
CGCTGAGGTGACTGGGCTGTCAGGGATCTTCTATACAGAACCTGGGTCACTATTT
GTAATAAAATGTTGGCTGGAGGCTGATCCAGTGTCTGACCTAAAAAAAAAAAAAAA
AAAAAAAAAAAGGGCCGCCGCACTGAGATGTCGACCTGAGTGGATAACAGGGTAATAAGC
TGGGCCCTATGCCAACCTGTTTATGGCAG

FIGURE 39

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913
><subunit 1 of 1, 730 aa, 1 stop
><MW: 78644, pI: 7.65, NX(S/T): 2
MGVCQRTRAPKEKSQLERAAALGPRKGGSGMFASGWNQTVPIEEAGSMAALLLPLLLLLPL
LLLKLHLWPQLRWLPAIDLAFAVRALCCKRALRARAIAAAAADPEGPEGGCSLAWRLAELAQ
RAAHTFLIHGSRRFSYSEAERESNRAARAFLRALGWDWGPDDGDSGEGSAGEGERAAPGAGD
AAAGSGAEFAGGDGAARGGGAAAPLSPGATVALLLPAGPPEFLWLWFGLAAGRRTAFVPTAL
RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLAELA
EVDPGPVPGYLLSPQSITDTCLYIFTSGTTGLPKAARI
SHLKILQCQGFYQLCGVHQEDVIYL
ALPLYHMSGSLLGIVGCMGIGATVULKSKFSAGQFWEDCQQHRTV
FQYIGELCRYLVNQPP
SKAERGHKVRLAVGSLRPTDWERFVRRFGLQLVLETYGLTEGN
VATINYTGQRGA
VGRASW
LYKHIFPFDSLIRYDVITGEP
PIRD
PQGHCMATSPGE
PGILLVAPVSQQSPFLGYAGGPELAQGK
LLKDVF
RPGDVFFNTGDL
LCDDQGFLRFHDRTGDTFRWKGEN
VATT
EV
FEALDFLQE
V
N
VYGVTVPGHEGRAGMA
ALVLRPPHALDLMQLYTHVSEN
NL
PPYARPRFLRLQESL
ATTETFK
QQKVRMANEGFD
PSTLSDPLVLDQAVGAYLPL
TARYSALLAGNLRI

Type II transmembrane domain:

amino acids 45-65

Other transmembrane domain:

amino acids 379-398

cAMP- and cGMP-dependent protein kinase phosphorylation site
starting at amino acid 136

CUB domain protein motif

amino acids 254-261

putative AMP-binding domain signature

amino acids 332-343

N-glycosylation sites

amino acids 37-40 and 483-486

FIGURE 40

CCGTGTTAACGTGAGGTTCCCCTAGATCTCGTATATCCCCAACACATACTCCACGCA
CACATCCCAAAGAACCTCGAGCTCACACAAACAGACACACCGCGCAGCTCCACCTTGCCGACACTCCGGC
GCTTGTCCATCTCCCTCCGGGGAGCCGGCGCGCTCCACCTTGCCGACACTCCGGC
GAGCCGAGGGCGCAGCCCTCAGGATTCTCGGCTCGAAGCTCGGATTCGAGCTCTGAACC
CCATGGTGGTTTAAACACTTCTTCTCTTCTCGTTGATTGACCCGGTTCCA
TCIGGGGCTAGAGGACAAGGCAGCAGGCCCTCCAGCCAGGCCCTGTTGGCTGCATCGT
CCATCTGGCTTATAAAAGTTGCTGAGCGCACTCCAGGGCTGCTCGTCCCTCGG
CTGGCAGAAGGGGGTGAACGAGGGCAGGGCAGGGCAGGGCAGGGCTGCTCTGGCAGGGCTTT
CGGCTTGAGGGCAAGGTGAAGAGGCCACGGCGCTGGGTTACCGAGCTGGATTGTATG
TTGCACCATGCCCTTGGATCGGGCTGTGATTCTTCCCCTTGGGCTGCTCTCCC
TCCCGCCGGGGCGGATGTGAGGCTCGAGCTCGGGAGGGTCCGCCAGGGCTACGGTGC
AAGGGATTCGCTGGCAGACCTCCCTACCGAGGATCGCAGGGAAACACTTAAGGAATCTG
TCCCTCAGGAATATACATCGTCAGGCCACAGAAATGGAGAACAGTTAACCCA
TCGAATTGAAAACCTTGTGAAGAGAACAGCCATTGTGCGCACCCTTGTGTCAGG
CATAGAAATTGACGAATTTCGGAGAGCTCTGGGAATGAGAACAGAAAAGTCA
TAAATGATGTTGACGGACATTAGGCATCTGTCATCGAGAAATTCA
TATGTTGACGGACATTAGGCATCTGTCATCGAGAAATTCA
TCACAGAGCTAAAAGGTACTACACTGGGTAATGTGATCTGGAGGAAATGCTCAATGAC
TTTGGGCTCGGCTCTGGAACCGGATTTCACTGGGATTAACCCCTCAGTAC
AGACTACCTGGGAATGTGAGGAAATACACTGACCAGCTCAAGGCACTTGGAGACGTGCCCC
TGAAAACTGAAGATTCAAGGTTACCCCGCCTTCATTGCTGCCAGGACCTTGT
ACTGTGGGAGAGAAGTGCACACCGAGTTCCAAGGTAGGCCAACCCAGGGTGTATCG
TGCCCTCATGAAGAGATGCTGACTGCCCATACTGCTGGGGCTTCCCAC
AAACACTATGCTCAAGCTCATGAAGGGCTCTGGCAATCAGGCTGACCTCGACACAGAG
TGGAACTGTGTTATAGATGCAATGCTCTGGGGCAGAGCAGTGGAGGGGCAATT
TGAGTCGGTCAATGGACCCGATAGATGCTGAAGGCAATTGAAACATGCAAGAAA
ACAGCATGCGAGGTGTCAGGAAAGGTTCTCAGGGATGTTGTCAGGCCAACCTGCTCAGCC
CTCAGATCTGGCGCTCAGCTCTGGAAAATTAAACACAGGTTCAAGGCCCTAC
GGAAAGACCAACAACCTGCTGAGGCCACAGCTGGACCGGCTGGTCA
AATTGAAGGCTCTCTAAAAGGTCTGGTCACTGCTGAGGATACACT
GTGACAGCGGGCACTGCTCAAGAGGAGGAATGCTGGAAAGGGCACAG
GCCTGAGATCATGAATGATGGGCTACCAACAGATCA
AAACATCCGAGGTGGATGTGGACA
TCACCTGGGCTGACACTTCTCATGACAGCAGATTATGGCTCTGGT
GAGGAGGGCTTCTGGTCAAGGAAATCTGCA
CAGAGGCCCGCAGTGGATCCCGACGGAGAGGGTGGACTCTTCT
CAGGCCAGCGTGGC
CACTCCCTGCTCTGGTCTCACCCTGCAITGTCTGGCAGAGACTGT
CTGGGTTTGGTCAAGGAAATCTGCA
ATCTGGGTTTGGTCAAGGAAATCTGCA
TTCTTACACTTGGACAATGGACCATGCCACAAA
ACTTACCGTTCTATGAGAAGAGAG
CAGTAATGCAATCTGCTCCCTTTTGTTTCCAAAGAGTACCGGGTGC
GAGACTGA
CTTCCTCTGGTCTCAGCTATCTGAGGAGGGACCTTGT
TTTCTAGAGAGAAATTCTACTCAA
ATTTTCGACCGAGGATTTCTACCTTCA
CTCACGGTGTAGGGTTTTTTCTCATTTAAAT

FIGURE 41

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914
><subunit 1 of 1, 555 aa, 1 stop
><MW: 62736, pI: 5.36, NX(S/T): 0
MPSWIGAVILPLLGLLLSPAGADVKARSCGEVRQAYGAKGFSLADI PYQEIA GEHLRICPQ
EYTCCTTEMEDKLSQQSKEFENLVEETSHFVRTTFVSRHKKFDEFFRELLENAEKSLNDMF
VRTYGM LYMQNSEVFQDLFTELKRYYTGGVNLEEMLNDFWARLLERMFQLINPQYHFSEDY
LECVSKYTDQLKPFGDVPRKLKIQVTRAFIAARTFVQGLTVGREVANRVSKVSPTPGCIRAL
MKMLYCPYCRLGLPTVRPCNNYCLNVMKGLANQADLDTEWNLFIDAMLLVAERLEGPFNIES
VMDPIDVKISEA IMMNMQENSMQVSAKVFQGCGQPKPAPALRSARSAPENFNTRFRPYNPEER
PTTAAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDESVTAGTSNEEECWNGHSKARYLPE
IMNDGLTNQINNPEVDVDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQDTSDESSGSG
SGSGCMDDVCPTEFVTTTEAPAVDPDRREVDSSAAQRGHSLLSWSLTCIVLALQRLCR
```

Signal peptide:
amino acids 1-23

FIGURE 42A

CCGACCGCTGGCGGAGCGCTGGCAAAAGAACCTGGAGTGCCAAGCTAAATAAGTTAGCTGAGAAAACGCAGC
CAGTTTCGACGCCCTGCGCCGGTGCCTAACAGCAAAGACCAAGCGGGCTCGCCGGACCGCCGCGGGG
TAGGGACCCGGTTGGCCCTCAGGCTCCCTAGCAGCGGGAAAAGGAATTGCTGCCGGAGTTCTCGGGAGGT
GGAGGGAGATCAGGAAACGGCTTCTCCTACTTCGCGCTGGTAGTGTGGGGAGATTGCAAACCCCTAGG
AAAGGACTGGGAAAATAGGCCCTGGAAAGTGGAGAAGGTGATCAGGAGGGCGGTCACACTACGGCAGTTATCTG
TCTGATCAGAGCCAGACGCGACCGCTCACCTCGCAGTTCTCCAGGTGTGGGGACCGCAGGACAGACGCCGA
TCCC CGCGCCCTCGTACAGCACTCCAGGAGAGTCAGCCTCGTCCCAAACGTGAGGGCGCTCGGCCACGA
AAAGTTCTGTCCTACTGTGATTTCTCAATTCTGCTGGTTTTCTCCAGAGAACTTTGGGTGGAGATATTA
ACTTTCTCTTTCTGGTGGAGGGCTGCTAGGGAGGGGGAGGAGGAGGAAAAGTGAATGTGC
TGGAGAAAGGGAGGGCCCTCTGGTCTCCGGAGTCCCATTAAGGCATCATTCTGGAAGATTAAGTGTG
CGGACATGGTGAAGCTGAGAGGAGGAGGATTCTTCCGGCAGGTGGAGAGTCTTACCGCTGTGGTGCATG
TGTGGCCCSCAGCGCCGGGGCGCTGGTCTCGCTGGAGTCTACCTGGGACCTGAGTGAATGGTCCCA
GGGGCTGTCTGGGGCATCCGCTTCCACAGGCTGTGTCTGTCTGGAAAGATCTAGCAATGGGG
CCCTGGCAGGATTCTGGATCTCTGGCTCTCACCTATGGTTCTCTGGCCAGGCCCTAGAAGAGGAGG
AAGAAGGGGCTTACTAGCTCAAGCTGGAGAGAAAAGTAGGCCAGCACAACTTCCACCTCCAGCCCCATCTCA
TTTCATCCCTAGCGGATGATCAGGGATTAGAGATGTGGTTACACGGATCTGAGATTAAACACTACTCTT
ACAAGCTCGTCCGAAGGGTTAAACTGGAGAACTATGTCCAGCTATTGCAACACATCCAGGAGTCAGT
TTTAACTGAAAGTACAGATAACACCGGACTTCACATCTATATAAGACCTACCCAAACCCAACTGTTTAC
CTCTGGACATGGCACCCCTACCTCAGAAACTGAGGGTGGATTCACGCACTGGTGGAAATGGCACT
TGGGTTTAACAGAAAAGATGCACTGGCCACAGGAAGAGGATTGATACCTTTGGTTCTGGGAAGTG
GGGATTACTATACAACTACAAATGTGACAGTCTGGGATCTGTGCTATGACTGTGATGAAACAGCAATGTG
CTTGGGACTATGACAATGGCATATACTCCACACAGATGCAACTCAGAGAGTACAGCAATCTTAGCTTCCATA
ACCCCAACAAAGCTATTTTATATACTGCCTATCAAGCTGTTCACTCACCCTGCAAGCTCTGGCAGGATT
TCGAACACTACCGATCCATTATACAACATAAACAGGAGAAAGATGTGCTCCATGCTTAGATGAAGCAA
TCAACAAACGTGACATTGGCTTAAAGACTTATGGTTCTATAACACAGCATTATCATTACTCTCAGATAATG
GTGGCCAGCTACGGCAGGGAGGAGTAACTGGCCTCTAGAGGTGAGCAAGGAACATATTGGGAAGGGAGTCC
GGGCTGTAGGCTTGTGCACTAGGCCACTCTGAAAAACAGGGACAGTGTGAGAAGGACTGTGCACTAC
ACTGGTACCCACTCTCATTTCTGCTGAAGGACAGATTGTGAGGACATTCAACTAGATGGCTATGATATCT
GGGAGGACATACTGGAGGCTCTGGCTCACCCCGAGTAGATATTGGCTATACTGGACCCCTATACCAAAGGC
AAAATGGCTCTGGGCGACAGGGCTATGGGATCTGGAAACACTGCAATCCAGTCAGCCATCAGAGTCAGCACTG
GAAATTGCTACAGGAAATCTGGCTACAGGACTGGGCTCCCCCTCAGTCAGCAACCTGGGACCGAACCG
GTGGCACAATGAACGGATCACCTGTCACTGGAAAAGTGTATGGCTTTCAACATCACAGCGACCCATATGA
GAGGGTGGACTATCTAACAGGTATCAGGAATCTGAAAGGCTCTACGGAGGCTCACAGTTCAACAAAAC
TGCAGTGCCGGTCAAGGTATCCCCAAGACCCAGAAGTAAACCTAGGCTCAATGGAGGGTCTGGGACCATG
GTATAAAGAGGAAACCAAGAAAAAGAAGCAAGAAAATCAGGCTGAGAAAAGCAAAGCAAAGCAAAAAAA
GAAGGAAGAAACAGCAGAAACAGTCTCAGGTAACACAGCAAATTGGCTGATATACTGCTGGCTTAAGCGTCA
GGCTTGTGTTCTAGCTGTGCACTCCAGGAGACTCTCCACCTGGCCGGCACACTGAAAAGTGTCTGCTCAGTG
CCAAGGTGCTACTCTGCAAGGCCACACTTGTAGAGAGTGGAGATTTATTCTCTGCTCCTTGAAGAAAACGTG
GTGAGCTCTGGTCACTGGCTCAGTCAGTCAGTCAGGACTGAGGAGTACAGGAGTACAGGAGTACAGGAGAACA
ACCTACCATCCCAAGCCATGCTAATTGAGGAAGTTACAGGGTAGCATGATTAACACTTGTATAAATTAC

FIGURE 42B

FIGURE 43

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296
><subunit 1 of 1, 515 aa, 1 stop
><MW: 56885, pI: 6.49, NX(S/T): 5
MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTGYLSWGQALEEEEEGALLAQA
GEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKPTLDKLAEGVKLENYYVQPICTP
SRSQFITGKYQIHTGLQHSIIIRPTQPNCPLDNATLPQKLKEVGYSTHMVGKWHLGFRNKEC
MPTRRGFDFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQ
ILASHNPTKPIFLYTAQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLDENINVTLA
LKTYGFYNNSSIIYSSDNGGPTAGGSNWPLRGSKGTWEGGIRAVGVHSPLLKNKGTVCK
ELVHITDWYPTLISLAEGQIDEEDIQLDGYDIWETISEGLRSRPRVDILHNIDPYTPRQKMAPG
QQAMGSGLQSSQPSECSTGNCLQEILATATGSPLSLSATWDRTGGTMNGSPCQLAKVYGF
TSQPTHMRGWTYLTGQES
```

Important Features:

Signal Peptide:

amino acids 1-37

Sulfatases signature 1.

amino acids 120-132

Sulfatases signature 2.

amino acids 168-177

Tyrosine kinase phosphorylation site.

amino acids 163-169

N-glycosylation sites.

amino acids 157-160, 306-309 and 318-321

TOBUTT*HHS82650

FIGURE 44

CGGACCGCTGGGTGCGAGTGGAGCGGAGGCCGAGCGGCTGAGGAGAGAGGAGGCCGG
TTAGCTGTCACGGGGTCCGGCCGGCCTCCGGAGGGGGCTCAGGAGGAGGAAGGAGGAC
CCGTGCGAGAATGCTCTGCCCTGGAGCCCTGGCTCCGCTGCTCTCTGGGTGGCAG
GTGGTTTGGGAACCGGGCAGTGCAGGATCACCGGGTTAGCATCGGCACGTAGCCT
GGGTCTGCACTATGGAACTAAACTGGCCTGCTACGGCTGGAGAACAGCAAGGG
AGTCTGAAAGCTACATCGCAACTGGATGTAAGTTGGTAGTGCTGGGACCAAACAAAT
GCAGATGCTTCCAGGATACACCGGAAACCTGCACTCAAGATGTGAATGAGTGTGGAATG
AAACCCCGGCATGCCAACACAGATGTGAATACACAGGAAGCTACAAGTGCTTTGCCT
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGAATCTAGGAATGTGCCATGATAA
ACTGTCAGTACAGCTGTGAAGACACAGAAGAAGGGCACAGTGCTGTGTCATCCTCAGGA
CTCCGCTGGCCCAAATGGAAGAGACTGCTAGATATTGATGAATGTGCTCTGGTAAAGT
CATCTGCTCCATACATCGAAGATGTGTGAACACATTGGAAAGCTACTACTGCAAATGTCACA
TTGGTTTGCAGAATCTGCAATATCATGGAGCATATGCTGATATGATAATAAATGAATGACT
ATGGATAGCCATACGTGCAGCCACCATGCCAATTGCTCAATACCCAAGGTCCTTCAGTG
TAAATGCAAGCAGGGATATAAGGCAATGGACTCTGGTCTGCTATCCTGAAAATCTG
TGAAGGAAGTCTCAGGACACTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGCTC
AAAACAGCATGAAAAAGAAGGCAAAATTTAAATGGTACCCAGAACCCACCAGGACTCC
TACCCCTAAGGTGAACTTGCAAGCCCTCAACTATGAAGAGATAGTTCCAGAGGCGGAAC
CTCATGGAGGATAAAAGGAAATGAAAGAGGAAATGAAAGAGGCTTGAGGATGAGAAAAGAG
AAGAGAAAGCCTGAAGAATGACATAGGGAGCGAACCTGCAAGGAGATGTTTCC
AAGGTGAATGAAGCAGGTGAATTGGCCTGATTCTGGTCAAAGGAAAGCGTAACCTCCAA
ACTGGAAACATAAAGATTAAATCTGGTTGACTGAGCTTCAATCATGGGATCTGTGACT
GGAAAAGCAGGATAGAGAATGATTGACTGGAAATCTGGCTGATCGAGATAATGCTATTGGC
TTCTATATGGCAGTCCGGGCTTGGCAGGTACAAGAAAGACATTGGCCTGAAACTCT
CCTACCTGACTCTGCAACCCCAAAGCAACTCTGGTTGCTCTGGTCAATTACCGGCTGGCCGAG
ACAAAGTGGGAAACCTTGAGTGGTGGAAAAACAGTAACAAATGGCCTGGCATGGGAG
ACCCAGACTGAGGATGAAAAGTGGAAAGACAGGAAATTCTGAGTCTATCAAGGAACTGATGC
TACCAAAAGCATCTTGAAGCAGAACGTGGCAAGGGAAACCCGGAAATCGCAGTGG
ATGGCGCTTGGCTTGGTTGAGGCTTATGTCAGAGTACGGCTTATCTGGGATGACTGAATG
TTACTATCTTATATTGACTTGTATGTCAGTTCCCTGGTTTTTGATATTGCTCATCATG
GACCTCTGGCATTAGAATTAAGATGCAATTGGCTTTAAATATCATACTGATCT
TAAGATGCTTCTTGATAAGATGCAATTGGCTTTAAATATCATACTGATCT
TCTCAGTCATTCTGAATCTTCCNCAATTATTAATNTGAAANGTCAGTTTATCTC
CCCTCCTCNGTATATCTGATTGTTGATANGTGTGNGCTCTCTACAAACATTCTA
GAAAATAGAAAAAAAGCACAGAGAAATGTTAAGTGTGACTCTTATGATACTCTTGG
AACTATGACATCAAAGATGACTTTGCTAAGTGGCTTAGCTGGCTTATGAGAAC
TTGTATATTAAATTCTTGATAATAAA

FIGURE 45

MPLPWSLALPLLSSWAGGFGNAASARHHGLLASARQPGVCHYGTKLACCYGWRNSKGVC
ATCEPGCKFGEVGPNKRCFPGYTGKTCSDVNNECGMKPRPCQHRCVNTHGSYKCFCLSGH
MLMPDATICVSRTCAMINCQSYCSEDTEEGPQCLCPSSGLRLAPNGRDLIDECASGKVICP
YNRRCVNTFGSYCKCHIGFELQYISGRYDCIDINECTMDSHCSHANCFNTQGSFKCKCK
QGYKGNGNLRCSAIPENSVKEVLRAPGTTIKDRIKKKLAHKNSMKKAKIKNVTPETRTPPK
VNLOPFNEYETIVSRGGNSHGKKKGNEEK

Signal peptide:

amino acids 1-21

EGF-like domain cysteine pattern signature.

amino acids 80-91

Calcium-binding EGF-like domains

amino acids 103-124, 230-251 and 185-206

FIGURE 46

GGGAGCTGCTGCTGTGGCTGCTGGTGTGCGCGCTGCTCCTGCTTGGTGACGCTGCTG
CGCTTCCTGAGGGCTGACGGCGACCTGACGCTACTATGGCCGAGTGGCAGGGACGACGCC
AGAATGGGAGCTGACTGATATGGTGGTGTGGGTGACTGGAGCCTCGAGTGGAAATTGGTAGG
AGCTGGCTTACCACTTGTCTAAACTAGGAGTTCTTGTGTGTCAGCCAGAAGAGTGCAT
GAGCTGGAAAGGGTAAAAGAAGATGCCTAGAGAATGCAATTAAAAGAAAAAGATATAACT
TGTGTTGCCCCCTGACCTGACCGACACTGGTCCCAGAAGCGGCTACCAAAGCTGTTCTCC
AGGAGTTGGTAGAACATCGACATTGGTCAACAATGGGAATGTCCAGCGTTCTGTGC
ATGGATACCAGCTTGGATGTCACAGAAAGCTAACAGAGCTTAACACTTACGGACGGTGTC
CTTGACAAAATGTGTTCTGCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA
ATAGCATCCTGGGTATCATCTGACCTCTTCCATTGGACTGTGCTAGCAAGCATGCT
CTCCGGGTTTTTTAATGCCCTCGAACAGAACATTGCCACATACCAAGGTATAATAGTTT
TAACATTGCCAGGACCTGTCAATCAAATATTGTGGAGAATTCCCTAGCTGGAGAAGTCA
CAAAGACTATAGGCAATAATGGAGACCAGTCCCACAAGATGACAACCAGTCGTTGTGCGG
CTGATGTTAACATCAGCATGCCAATGATTGAAAGAAGTTGGATCTCAGAACACCTTTCTT
GTTAGTAACATATTGTGCAATACATGCCAACCTGGCTGGATGCAAGACTCTTCTTATTAAAATC
GGAAGAAAAGATTGAGAACCTTAAGAGTGGTGTGGATGCAAGACTCTTCTTATTAAAATC
TTTAAGACAAAACTGAAGAGCACCTGTACTTTCAAGCCACTGGAGGGAGAAATG
GAAAACATGAAAACAGCAATCTCTTATGCTCTGAATAATCAAAGACTAATTGTGATT
ACTTTTAATAGATATGACTTGTCTCAACATGGAATGAAATAAAAATAATAAAAAG
ATTGCCATGAATCTGAAAA

FIGURE 47

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343
><subunit 1 of 1, 289 aa, 1 stop
><MW: 32268, pI: 9.21, NX(S/T): 0
MVVWVTGASSGIGEELAYQLSKLGVSLLSARRVHELERVKRRCLENGNLKEKDILVLPLDL
TDTGSHEAATKAVLQEFGRIDILVNNGGMSQRSLCMDTSLDVYRKLIELNYLGTVSLTKCVL
PHMIERKQGKIVTVNSILGIIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIVSNICPGP
VQSNIVENSLAGEVTKTIGNNGDQSHKMTTSRCVRLMLISMANDLKEVWIEQPFLLVTLW
QYMPTWAWWI TNKMGKKRIENFKSGVDADSSYFKIFKTKHD
```

Important Features:

Signal Peptide:

amino acids 1-31

Transmembrane domain:

amino acids 136-157

Tyrosine kinase phosphorylation site.

106-113 and 107-114

Homologous region to Short-chain alcohol dehydrogenase

amino acids 80-90, 131-168, 1-13 and 176-185

FIGURE 48

GGCAGCTGGGCACGCCATCAGCTGTCGGCGCTCTCTCCAGGTGGGCAGGGTTTC
GGCTGGTGAGCATGTCGGTGGACAGCACCATCTCAATCAACAGCATATTCCG
TTCGATCTTCAACACTAACAGTATTGTTAGGTCCTCGCGCACACCTGGGCTCTGTC
TGATGCTGCTGAGCTCCCTGGTGTCTCGCTGGTTCTGTCACCTGGCTGGATCTGTT
TTCGTCCTATGATTTCGATTGTTGATCACCCACATGTCATACAGTCGAGCTGAT
GTGGCTCAGTTCCGGAAGGTCAGGAAAGCCCAGGGCAAGGCTAAGGACACTGAGCCTCA
ACCAAGGCCAGCTGACCTCATGCTTGTCTTCAGGCGCTCAGCGTCCCTGTC
GACAGCGTGCCCCGGCCCCAGCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGGCC
TGGCAGGGCAGTGTGAGGAGGCCACATCTGCGAGCGCTCCCTGGTGGAGCAC
CTGGGCTCTACTGCTGCCACTGCTTGTAAAAGGCAGCAGCAACAGAAGTCAATTCTGGT
CAGTGGCTGGGGTCTCTGCGAGCTGAGGGACTCAGGCCCTGGGGCGAAGAGGTTGGGGT
GCTGCCCTGAGTTGCCAAGGCCATAACCAACTACAGCCAGGGCTAGACCTGGGCCCTGCT
GCAGCTGCCACCCAGGCCACACCCCTCTGCTGGCCAGGCCCTCATGCTTCC
CTTTGGAGCTCTCTGCTGGGCCACTGGCTGGGATCAGGACACAGTGTCTCTGGAGC
CTACGCAATCTGCGCTGGTCTCATCAGTCGCCACATGTAACGTATCTACAACCAGCT
GCACAGCAGACACTGTCAACCCGGCCGGCTGGAGTGTATGGGGGCCCCAGCTG
GGGTGAGGGCCCCCTGTCAAGGGAGATTGGGGGCCCTGTGCTGTGCTGGAGCTCAGGVA
CACTGGGTTAGGCTGGCATCATCAGCTTGTCAAGCTGTCAGGCCAGGAGCCTCTGT
GCTGCTGACCAACACAGCTGTCACAGTCTCTGGCTGAGGCTCGAGGTTCTGGGGCAGCTT
TCTCTGGCCAGAGGCCAGAGACCCGGAGATGAGTGTAGGAGACACTGTGTTAGCTGTGGA
TCTCTGGAGGACAGCAGGGCTTCCAGGGCAGGAGCACCTCCCATGGCCCTGGAGGGCAGGCT
GATGCACCAAGGGACAGCTGGCTGTGGGGAGGCCCTGGTGTAGAGGAGGGCTGTGCTAAGT
CTGCCACTGCTTCACTGGGCCAGGGCCCAGAGGAATGGAGCTGAGGGCTGGGGACAGA
CCGGAGGAGTGGGGCTGAAAGCAGCTCATCTGCTGGAGGCCCTACACCCACCTGTGGGGGG
CTACGACATGGGCCCTCTGGCTGGGCCAGCCTGTCACACTGGGAGGCCAGCTGGGGCCC
TCTGCTGCCCTATCTGACCCACCTGCTGCTGATGGGGAGCTGGCTGGGTTCTGGGACGG
GCCCGCCAGGAGCAGGACATCAGCTCCCTCCAGACATGCCCCGTGAGCCTCTGGGGCT
GGCTCGACCGGGCTCATGCTGAGCTCTGGGGTGTAGCAGGCCCTTATCTGGGGGATGG
TGTGTACCACTGCTGTGGGTGAGCTGCCAGCTGTGAGGGCCGTGCTGGGGCACCACACTGGT
CATGAGGTGAGGGACACATGGTTCTGGGCCGGCTGCAAGCTGGAGATGCTTGGCAAGG
CCCGCCAGGGCCGGCTTCTACCGCGCTCTGGCTCATGAGGACTGGGTCTGAGCTTGG
ACTGGCAAGTCTACTTCTGGCAGGAACAGAGGCCAGGGCTGAGCTGGAGCTGGCC
AACATAAGGCAACCAACAGCTGTCAGGAGGGACCTGGGCACTTCTCAGGACAAGAAGATG
AGGCAGGCAATGGGCAATTACTGGGCTCTGGGCTCCACCTCTGGTGTGATTCTCAGGAC
CAGGGCAGGGCCAGAGGCCAGCAGCTGGAGGAGAACCTGGCTGGGACACCCAACTCTG
AGGCCTGTCAGCTTCTCCCTCTTACTCTTCAAGATAACATCAGGCCAGCCAGCTGG
TAAAAATTCTTTTTGGGGCAGCAGCTTCTTTTAAACTTAAATAATTGTTAC
AAAAA

FIGURE 49

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571
MLLSSLVSLAGSVYLAWLFFVLYDFCIVCITYAINVSLMWLSFRKVQEPQGKAKRHGNTV
PGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEAAATELNSWSVVLGSLQREGLSPGA
EEVGVAALQLPRAYNHYSQGSDIALLQLAHPTTHTPLCPQPAHRFPFGASCWATGWDQDTS
DAPGTLRNLRLRLLISRPTCNCIYNQLHQRHLSNPAPGMLCGGPQPGVQGPCQGDGGPVLC
LEPDGHWVQAGIISFASSCAQEDAPVLLNTAAHSSWLQARVQGAAFLAQSPETPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQQLACGGALVSEEAVLAAHCFIGRQAPEEWSV
GLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLAQPVTLGA SLRPLCLPYPDHHLPDGERG
WVLGRARPAGISSLQTVPVTLLGPRACSRLHAAPGGDGSPILPGMVCTS AVGELPSCEGLS
GAPLVHEVRGFTWFLAGLHSFGDACQGPARPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAEP
GSCLANISQPTSC
```

Important features:

Signal peptide:

amino acids 1-15

Homologous region to Serine proteases, trypsin family

amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.

amino acids 37-40 and 564-567

Kringle domains

amino acids 79-96, 343-360 and 235-247

19970614.101501

FIGURE 50

CGGGCCGCCCCGGGCCCCATTGGGCCGGGCCTCGTCGGCGGGCAGTGAGGCCAGGCTGG
GCCGCGTCCCTGAGTCCCAGTCGGCGCGCGCGCAGGGCAGCCTTCACACGGGGAG
CCCAGCTGTCAAGCGCCTCACAGGAAGATGCTGCGTCGGCGGGCAGCCCTGGCATGGGTGT
GCATGTGGGTGAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG
TCCCTGAAGACCCAGTGGTGGCACTGGTGGCACCGATGCCACCCCTGTGCTGCTCCTTC
CCTGAGCCTGGCTTCAGCCTGGCACAGCTAACCTCATCTGGCAGCTGACAGATAACAAACA
GCTGGTGACAGCTTGCTGAGGGCCAGGGCAGCAGCCTATGCCAACCGCACGGCCC
TCTTCCCGGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGCTGCGTGGCG
GACGAGGGCAGCTTCACCTGCTTCGTGAGCATCCGGGATTCGGCAGCGCTGCCGTAGCCT
GCAGGTGGCCGCTCCCTACTCGAAGCCCAGCATGACCCCTGGAGGCCAACAGGACCTGCGC
CAGGGGACACGGTGACCATCAGTGTCCAGCTACAGGGCTACCTGAGGCTGAGGTGTT
TGGCAGGATGGCAGGGTGTGCCCCCTGACTGGCACAGTGCACAGCTGCGAGATGGCCAACGA
GCAGGGCTTGTGATGTGACAGCGTCTGGCGGGTGGTGTGGGTGCGAATGGCACCTACA
GCTGCCTGGTGCGCAACCCCGTGTGCAAGCAGGATGCCACRGCTCTGTACCATCACAGGG
CAGCCTATGACATTCCCCCAGAGGCCCTGGGTGACCGTGGGGCTGTCTGTCTCAT
TGCACTGCTGGTGGCCCTGGCTTCTGTTGCTGGAGAAAGATCAAACAGAGCTGTGAGGAGG
AGAATGCAGGAGCTGAGGACCAGGATGGGGAGGAAGGCTCAAGACAGCCCTGAGCCT
CTGAAACACTCTGACAGCAAAGAAGATGATGGACAAGAAAATAGCTTGACCATGAGGACAGG
GAGCTGCTACCCCTCCCTACAGCTCTACCCCTGGCTGCAATGGGGCTGACTGTGAGCCC
TGCCCCCAACAGATGCATCTGCTCTGACAGGTGGCTCTCTCCAAAGGATGCGATAACAC
AGACCACTGTCAGCCTTATTCCTCAATGGACATGATTCCAAGTCATCTGCTGCCCTTT
TTCTTATAGACACAAATGAAACAGACCAACACCTTAGTTCTCTAAGTCATCTGCCCTG
GCCCTTATTCACAGTACATACATTCTTAGGGACACAGTACACTGACCAACATCACCACCC
TTCTTCCAGTGTGCGTGGACCATCTGGCTGCCTTTCTCCAAAGATGCAATATTCA
CTGACTGACCCCTGCCTTATTCACAAAGACACGATGCACTAGTCACCCGGCTTGT
TCCAATGGCCGTATACTAGTGTACATGTTGAGCTGCTGGAGTCTAGAAGCTGTTCT
TCTTCTCAGACAGGGACAGTGCAGCCCTCAACATCTCCTGGAGTCTAGAAGCTGTTCT
CCCTCCCTCCCTGCCCAAGTGAAGACAGGGCAGGGCAGGAATGCTTGGGACACCG
AGGGGACTGCCCTCCACCCCAACATGGTGCTATTCTGGGGCTGGGGCAGTCTTCT
TTGCCCTCTGGCAGCTCTGGCCTCTGGTAGAGTGAGACTTCAGACGTTCTGATGCCT
GATGTCATCTCCCTGCCCAAGGAATGGAAGATGTGAGGACTCTAATTAAATGTGGGAC
TCGGAGGGATTGTAAACTGGGGTATTTGGGAAAATAATGCTTTGTAAAAAAA
AAAAAAAAAAAAAA

FIGURE 51

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386
><subunit 1 of 1, 316 aa, 1 stop, 1 unknown
><MW: -1, pI: 4.62, NX(S/T): 4
MLRRRGSPGMGVHVGAAALGALWFCLTGALEVQVPEDPVVVALVGTDATLCCSFSPPEPGFSLAQ
LNLIWQLTDTKQLVHSFAEGQDQGSAYANRTALFPDLIAQGNASLRLQRVRVADEGSFTCFV
SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDTVTITCSSYQGYPEAEVFWQDGQGVPL
TGNNVTTSQMANEQGLFDVHSVLRVVLGANGLTYSCLVNPVLQQDAHXSVTITGQPMTFPPEA
LWVTVGSLSVCLIAALLVALAFVCWRKIKQSCEENAGAEDQDGEGEGSKTALQPLKHSDSKED
DGQEIA
```

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 251-270

N-glycosylation site.

amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC

amino acids 217-234

09978541-101501

FIGURE 52

TTCTGTGACCCCTTGAGAAAAAGAGTGGTGTAAATGTGCCACGTCTCTAAGAAGGGGGAGTC
CTGAACTTGTCTGAAGCCCTGTCGCTAAGCCTGAACTACGTTCTTAATCTGAAGTCG
AGGGACCTTCGCTCTTGTAGGGACTCTTCTCTGCTTCAGCAACATGAGGCTTTCT
TGTGGAACGCGGTCTGACTCTGTCCTTGATTGGGCTTGTATCCCTGAAACCA
GAAGTGAAAATTGAGGTTCTCAGAACAGGCAATTCTCATCTGCCATCGCAAGACCAAAAGGAGGGGA
TTTGATGTTGCTCACTATGAGGCTACTTAGAAAAGGACGGCTCTTACTTCACTCCACTC
ACAAACATAACAATGGTCAGCCATTGGTTACCTGGGCATCTGGAGGCTCTCAAAGGT
TGGGACCAAGGGCTTGAAGGAAGGAAAGGTTGTAGGAGAGAAAGAGAAACTCATCATTCCCTGC
TCTGGCTATGGAAAAGAAGGAAAGGTTAAATTCCCCAGAAAGTACACTGATATTAAATA
TTGATCTCTGGAGATCGAACATGCCATGAATCAITCCAAGAAATGGATCTT
AATGATGACTGAAACTCTAAAGATGAGGTTAAAGCATATTAAAGAAGGACTTGTGAAA
ACATGGTGGCTGTGAAGTAAAGTCATCATGATGCTTGGGGAGGATATTGATAAG
AAAGATGAAGACAAAGATGGTTATATCTGCCAGAGAATTACATATAACAGATGAGTTA
TAGAGATACTACCCCTTTAATATAGCACTCATCTTCAAGAGAGGGCAGTCATCTTAA
AGAACATTTTATTTTATACATGTTCTTGTCTTGTCTTATTTTATATTTTATTTT
CTGACTCTATTTAAAGAACCCCTTGTGTTCTAAGTACCATTTCTGTGATAAGTATT
GGGAAGAAAAGCTAATTGCTTGTGAAGACTCTGCCAAATTTCACCTCACAG
ATATGAAGCTTGTCTTACTTCTCACTTAAATTAATGTTGCAACTGGGAATATACC
ACGACATGAGACCAAGGTTAGCACAATTAGCACCCCTATTTCTGCTTCCCTATTTCT
TCCAAGTTAGAGGTCAACATTGAAAGGCTTTGCAATAGGCCAAAGGCTTGTCTTATTTCT
GTATAATGAAGATAGTTATGTGTAATGGCTGAGTCTGTGCTTGAGGACAGAGGAAA
TGGTTGTTGACCTGACTTGTAAATGGCTACTTACTAAGGAGATGTGCAATGCTGAAG
TTAGAAAACAGGTAAATGCCAGGATGGTGGCTCATGCTGTAAATCCAGCACTTGGGAG
GCTGAGGGGGGGATCACCTGGAGTTGGAGCTGGAGACCCAGCCTGACCAACAGGAGAAA
CCCTATCTCTACTAAAATACAAAGTAGCCCGGCTGGTATGCGTGCTGTAAATCCAGCT
ACCCAGGAAGGCTGAGGCCAGAACATCTGAACCCGAGGGCGAGGTTGCGGTAAAGCCGAG
ATCACCTNCAGCTGGACACTCTGTCGAAAAAGAAAAGAACACCGTTAATACCATATNA
ATATGATGATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTGGCTCTAGTGAT
TGGTGCCCTATTATGATAAAATAGGACAAATCATTATGTTGAGTTCTTGTAAATAAAATG
TATCAATATGTTATGAGGAGTAAAGGTTATTTATATTCAATATTACTCTTAAAGGC
TAGCGGAATATCCTTCTGGTTCTTAAATGGTAGTCTAGTATATTACTACAATAACA
TTGTATCATAAGATAAAAGTAGTAAACCACTCATTTCCCATTTCTGCTCATCAAAAAC
TGAAGTTAGCTGGGTGTGGTGGCTCATGCCGTAAATCCAGCACTTGGGGGCAAGGAGGG
TGGATCACTTGGAGATCAGGAGTTCAAGACCAAGCCTGCCAACATGGTAAACCTTGTCTCA
CTAAAATACAAAATAGCCAGGCGTGGTGGTGCACACCTGTAGTCCCAGCTACTCGGAG
GCTGAGACAGGGAGATTGGCTGAACCCGGAGGGTTGCAAGTGGAGGTTGAGCTGAGC
ATGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTAAAAAAAGAACAGCAGA
CCTACAGCAGCTACTATTGAAATAACTATCCTGGATT

FIGURE 53

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44194
><subunit 1 of 1, 211 aa, 1 stop
><MW: 24172, pI: 5.99, NX(S/T): 1
MRLFLWNAVLTLFVTSILGALIPEPEVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKD GSL
FHSTHKHNNGQPIWFTLGLILEALKGWDQGLKGMCVG EKRKLIIIPPA LGYGKEGKGKIPPEST
LIFNIDLLEIRNGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKHGAVVNESHHDALVED
IFDKEDEDKDGFISAREFTYKHDEL
```

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 176-179

Casein kinase II phosphorylation site.

amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.

amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 78-114 and 118-131

EF-hand calcium-binding domain.

amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain

amino acids 183-203

FIGURE 54

AATAAAAGCTTCCCTTAATGTTGTATATGTCCTTGAGATCACCGTGCACTTTTTTTAGCAT
CCAACCATTCTCCCTTGAGTTCCTGCCCTCAAATCACCTCTCCGTAGCCCCACCGA
CTAACATCTCAGTCTCTGAAAATGCACAGAGATGCCCTGGCTACCTCGCCCTGCCCTCAGCT
CACGGGGCTCAGTCTCTTTCTCTTGGTGCACAGAGCAGGAGCATGGAGGTACAGTAC
CTGCCACCCCTCAACGTCCTCAATGGCTCTGACGCCCGCTGCCCTGCCACCTTCACCTCTGC
TACACAGTGAACCAAACAGTCTCCCTGAACATGGACTTACCAAGGAGTGAACAAACTGCTC
TGAGGAGATGTTCTCCAGTCCAGTGAAGATCATTAACCTGAAGCTGGAGCGGTTCAAG
ACCGCGTGGAGTTCTCAGGGAACCCCAAGCTAGATGTCTGGTGTAGAAGAACAGTG
CAGCCGGAGGATGAGGGATTATCAACTGCTACATCATGAAACCCCCCTGACGCCACCGTGG
CCATGGCAAGATCATCTGCAAGCTCATGGAAAGAGCCCCCTGACGCCACCGTGG
CCGTGATTGGGGTGCTCCGTCGGGGCTTCTGGCTGTGGTCATCTGGTGTGATGGTG
GTCAAGTGTGAGGAGAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACCGAGGA
GGAGGGCAAGGACGGCTGAAGGCCACCCGGATGATGCCCAAGTAGTGGGTGGCGCC
CTGCAGCCTCCGTGCTCCGTCTCTCCCTCTCCGCGCTGTACAGTGAACCTGCCCTGCTCG
CTCTGGTGTGCTTCCGTGACCTAGGCCACGGGGCACCTGGGGCTCTGAAACCCCG
ACTTCGTATCTCCCACCCGACCAAGAGTGAACCACTCTTCCATCCGAGAAAACCTGCCA
TGCTCTGGGACGTTGCGGCCCTGGGGAGAGGAGAGAAAAGGGCTCCACCTGCCAGTCCCTGG
GGGGAGGCAGGAGGCACATGTGAGGGTCCCCAGAGAGAAGGGAGTGGTGCGAGGGTAGA
GGAGGGGGCCGTGACCTGGCCAGTGGCTGCCAGTGGCTTCAGAGAGGACCTGGTGG
GGAGGGAGGGCTTCTGTGCTGACAGCGCTCCCTCAGGGGGCTTGGCCTGGCACGGCTG
TGCTCTCCCTCTGCTCCAGCCAGAGCAGCCATCAGGCTGGAGGTGACGATGAGTCTGTA
AACTTGGAGGGCATGTTAAAGGGATGACTGTGCACTCCAGGGACTGACGGAAAGGCCAGG
CTGCAGGCAAGGACATGTGCCCTGGCCAGGGAGGGCATGTGGGCCCTCGTTCCATT
GCTAGTGGCCCTCTGGGGCTCTGGCTTAATCCTTAGGACTGTGGATGAGGCCAG
ACTGGAAGAGCAGCTCCAGTGAAGGGGCATGTTTCCAGCGGGGACCCACAAACAGAGGCC
AGTTTCAAAAGTCAGCTGAGGGGTGAGGGGTGGGGCTCATGGTAATGCAAGGTTGCTGAG
GCTCTGCCCTCTCATGGGTAACCCCTGCCCTGGCAGGGGAGCCAAGGCTGGGAAT
GAGGAGGGCATGACAGGGTGGGGCAGCTTCTTGGGCTTCAGTGAGAACTCTCCAGTT
GCCCTTGGGGGTTTCCACCTGGCTTGGCTACAGAGGGGAAGGAAAGCCTGAGGCCG
GCTAAAGGGGAGGGCCTGGAAACCTGAGCTGCCAATGCCAGCCCTGCTCCATCTGCCAGG
CTACTCGCTCTCTCCAAACAACCTCCCTGCTGGGGACAAAGTGAACATTGAGGCCAGGC
ACAGTGGCTCACGCCGTATACTCCAGCACTTGGGAGGCCAGGCGGTGGATTACCTCCAT
CTGTTAGAAGAATGGGAAAACCCCATCTCTACTAAAATACAAGAATTAGCTGGCGTG
GTGGCGTGTGCTGTAAATCCAGCTATTGGGAGGTGAGGCAGGAGAATCGCTTGAGCCG
GGAAGCAGAGGTTGCACTGAACATGAGATAGTGAATGCCACTGCAATTAGCCTGGGTGAC
ATAGAGAGACTCCATCTCAAAAAAA

FIGURE 55

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415
<subunit 1 of 1, 215 aa, 1 stop
<MW: 24326, pI: 6.32, NX(S/T): 4
MHRDAWLPRPAFSLTGLSLLFFSLVPPGRMSMEVTVPATLNVNGSDARLPCTFNSCYTVNHKQ
PSLNLWTYQECCNNCEEMFLQFRMKIINLKLERFQDRVEFGNPSKYDVSVMLRNQVPEDEGI
YNCYIMNPDRHRGKIHQLVLMEEPPERDSTVAVIVGASVGGFLAVVILVLMVVVKCVRK
KEOKLSTDLLKTEESEGKTGEGNPDDGAK
```

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 161-179

Immunoglobulin-like fold:

amino acids 83-137

N-glycosylation sites.

amino acids 42-45, 66-69 and 74-77

FIGURE 56

GTTGTATATGCTCTGAAGTACATCCGTGATTTTTTAGCATCCAACCACCTCCCTTGTA
GTTCTCGCCCCCTCAAATCACCTCTCCCTTAGGCCACCCNACTAACATCTCAGTCCTCTGAA
AATGCACAGAGATGCTGGCTACCTCGCCCTGCCTTCAGCCTCACGGGCTCAGTCCTCTT
TCTCTTGGTGCCCCCAGGAGCGGAGCATGGAGGTACAGTACCTGNCCACCCCTCAACGTCC
TCAATGGCTCTGACGCCCGCTGCCCTGCCCTCAACTCTGCTACACAGTGAACCACAAAC
AGTTCTCCCTGAACTGGACTTACCCAGGAGTGCACAAACTGCTCTGAGGGAGATGTTCTCCAG
TTCCGCGATGAAGAGTCAATTAACTGAGGTGGAGCGGTTCAAGACCCGCTGGAGTTCTCAGG
GAACCCCCAGCAAGTACGATGTGTCGGTGTGCTGAGAACGTGCAGCCGGAGGATGAGGGGA
TTTACAACCTGCTACATCATGAACCCCCC

FIGURE 57

TCACGGGGCTCATCTTTCTTTGGTCCCCACCAAGGACGGAGCATGGAGGTNCACATA
CCTGCCACCCCTCAACGCTCTCAATGGCTTGACGCCCCCTGCCTGCACCTTCAACTCCNG
CTACACAGTGAACCACAAAAGCTTCTCCCTGAACTGGATTACCAAGGAGTGCACAACTGGC
TCTGAGGAGATGTTCTCCAGTCCCCGATGGAAGATCATTTAACCTGAAAGCTGGAAGCGG
TTTTCAAGAACCGCGTGGAAAGTTCTCAGGGAACCCAGCAAGTACGATGTGTCGGTGTGATGC
TGAGAACAGTGCAGCCGGAGGATGAGGGGATTTACAACGTACATCATGAACCCCCC

FIGURE 58

FIGURE 59

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44189
><subunit 1 of 1, 412 aa, 1 stop
><MW: 46658, pI: 6.65, NX(S/T): 4
MGLHLRPYRGVLLPDGLLFLLLLMLLADPALPAGRHPPVLVPGDLGNQLEAKLDKPTVVH
YLCSSKKTESYFTIWLNLLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGVDVRVPFGKTFSL
EFLDPSKSSVGSYFHTMVESLWGTYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEEMYQ
LYGGPVVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNNRI
PVIGPLKIREQQRSAVSTSLLPYNTWSPEKVFVQTPQTINYTLRDYRKFFQDIGFEDGWLM
RQDTEGLVEATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGTVNLSALQCQ
AQWQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLLGP
```

Important features:

Signal peptide:

amino acids 1-28

Potential lipid substrate binding site:

amino acids 147-164

N-glycosylation sites.

amino acids 99-102, 273-276, 289-292 and 398-401

Lipases, serine proteins

amino acids 189-201

Beta-transducin family Trp-Asp repeat

amino acids 353-365

01-1601-1458758-01

FIGURE 60

CGGACGCGTGGCGGACGCGTGGGCGGCCAGCGGGCGCAGCGGCAGCGCACATGGAGAGCGGG
GCCTACGGCGGGCCAAGCGGGCGCTCCCTGACCTGCGGCGTTCTGACGCAGCCGA
GGTGGTGGCGCGCGCGTGTGCTTGGTCCTGCCTTGATCGTGTTCCTGCATCTATGGT
AGGGCTACAGCAATGCCAACGAGTCTAACGAGATGTACTGCGTGTCAACCGCAACGAGGAT
GCCTGCGCTATGGCAGTGCCATCGGGGCGCTGGCCTTCCTGGCCTCGGCCTCTTCTGGT
GGTCGACGCGTATTCCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTATTGGT
ACCTGCTCTCTCAGCTCTGGACCTTCTGTGGTTGGTTCTGCTTCCTCACCAAC
CAGTGGGCAGTCACCAACCCGAAGGACGTGCTGGTGGGGCGACTCTGTGAGGGCAGCCAT
CACCTTCAGCTTCTTCCATCTCTGGGTGTGCTGGCCTCCCTGGCCTACCAGCGT
ACAAGGCTGGCGTGGACGACTTCATCCAGAATTACGTTGACCCCACCTGGACCCAAACACT
GCCTACGCTCTCACCCAGGTGCATCTGTGACAACACTACCAACAGCCACCCCTCACCCAGAA
CGCGGAGACCAACCGAGGGCTACCGCCGCCCCCTGTGTACTTGAGTGGCGGTTAGCGTGGAA
GGGGGACAGAGAGGGCCCTCCCTCTGCCCTGGACTTCCCATCAGCCTCCTGGAAACTGCCA
GCCCTCTCTTACCTGTCCATCTGTGCACTGACACACAGCTAAGGAGGCCATAGCC
TGGGGGGCTGGCAGGCCACACCCAAAGTGCCTGTGCCAGAGGGCTTCAGTCAGCCGCT
CACTCCTCCAGGGCCTTTAGAAAGGGTTTTAGCTAGTGTCTTCCTCGCTTAATGA
CCTCAGCCCCGCTGCACTGGCTAGAAGCCAGCAGGTGCCATGTGCTACTGACAAGTGCCT
CAGCTCCCCCGGCCGGGTCAAGCCGTGGGAGGCCGCTATTATCTGCGTTCTGCCAAG
ACTCGTGGGGGCCATCACACCTGCCCTGTGCACTGGGAGCCGACCAGGCTCTGTGCT
CTCAGGTTGCTCCCTGTGCCACTGCTGTATGATCTGGGGGCCACCACCCCTGTGCCGGT
GGCCTCTGGCTGCCCTGGGTGTGAGGGCGGGGCTGGTGTCTAGGCACCTCCCTGCT
CTCCCACCCCTGGCAGCAGGGAAAGGGTTGCTGACAACACCCAGCTTATGTAATATT
TGCAAGTTGTTACTTAGGAAGCCTGGGAGGGCAGGGGTGCCCATGGCTCCAGACTCTGTC
TGTGCCAGTGTATTATAAAATCGTGGGGAGATGCCGGCTGGATGCTGTTGGAGACG
GAATAATGTTCTCATTCAAAG

FIGURE 61

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48304
<subunit 1 of 1, 224 aa, 1 stop
<MW: 24810, pI: 4.75, NX(S/T): 1
MESGAYGAAKAGGSFDLRRFLTQPQVVARAVCLVFALIVFSCIYGEGRSNAHESKQMYCVFN
RNEDACRYGSAIGVLAFLASAFFLVVDAYFPQISNATDRKYLVIDLLLFSALWTFLWFVGFC
FLTNQWAVTNPKDVLVGADSVRAAITFSFFSIFSWSGVILASLAYQRYKAGVDDFIQNYVDPTP
DPNTAYASYPGASVDNYQQPPFTQNAETTEGYQPPPVV
```

Important features:

Type II Transmembrane domain:

amino acids 1-45

Other transmembrane domains:

amino acids 74-90, 108-126 and 145-161

N-glycosylation site.

amino acids 97-100

FIGURE 62

FIGURE 63

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49152
><subunit 1 of 1, 802 aa, 1 stop
><MW: 88846, pI: 6.41, NX(S/T): 7
MPVVAEAPQVAGGQQGDGGEEAEPEGMFKACEDSKRKARGYLRLVPLFVLLALLVLASAGVL
LWYFLGYKAEVMSVQVYSGSLRVLNRFHSQDLTRRESSAFRSETAKAQKMLKEKITSTRLGT
YYNSSSSVYSFGEGPLTCFFWFILQIPEHRRMLMSPEVVQALLVEELLSTVNSSAAVPYRAEY
EVDPEGLVIILEASVKDIAALNSTLGCGYRSYVGQGVRLKGPDHLASSCLWHLQGPKDILML
KLRLEWTIACERDRILAMYDVAGPLEKRLLITSVYGCQRQEPVVVEVLASGAIMAVVWKGLHSY
YDPFVLSVQPVVVFQACEVNLTLDNRQLDSQGVLSPTPYFPSSYSPQTHCSWHLTVPSLDYGLAL
WFDAYALRRQKYDLPCTQGWTTIQNRRLCGLRILQPYAERIPVVATAGITINFTSQISLTGP
GVRVHYGLYNQSDPCPCGEFLCSVNGLCVPACDGVKDCPNCGLDERNCVRATFQCKEDSTCIS
LPKVCDCGQPDCLNGSDEEQCQEGVPGTFTFQCEDRSCVKKPNPQCDGRPDRCRDGSDEEHCD
CGLQGPSSRIVGGAVSSEGEWPWQASLQVRGRHICGGALIADRWTAAHCFQEDSMASTVL
WTVFLGKVWQNSRWPGEVSFKVSRLLLHPYHEEDSHDYDVALLQLDHPPVRSAAVRPVCLPA
RSHFFEPGLHCWITGWGALREGGPISNALQKVDVQLIPQDLCSEAYRYQVTPRMLCAGYRKG
KKDACQGDGGPLVCKALSGRWFLAGLVSWGLGCGRPNYFGVYTRITGVISWIQQVTT

Important features:

Type II transmembrane domain:

amino acids 46-67

Serine proteases, trypsin family, histidine active site.

amino acids 604-609

N-glycosylation sites.

amino acids 127-130, 175-178, 207-210, 329-332, 424-427, 444-447

and 509-512

Kringle domains.

amino acids 746-758 and 592-609

Homologous region to Kallikrein Light Chain:

amino acids 568-779

Homologous region to Low-density lipoprotein receptor:

amino acids 451-567

FIGURE 64

GCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGTGTGGCTTGGCATCCTGCAGCCC
TACGCCAGAGGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTCACCTCCAGAT
CTCCCTCACCGGGCCCGGTGTGGGTGACTATGGTTGTACAACCAGTCGGACCCCTGCC
CTGGAGAGTTCCTCTGTTCTGTGAATGGACTCTGTGTCCCTGCCGTGATGGGTCAAGGAC
TGCCCCAACGGCCTGGATGAGAGAAACTCGCTTGCAGAGGCCACATCCAGTGCAGGAAAGAGGA
CAGCACATGCATCTCACTGCCAAGGTCTGTATGGCAGCCTGATTGTCTAACGGCAGCG
ATGAAGAGCAGTGCAGGAAGGGTGCCATGTGGACATTCACCTCAGTGTGAGGACCCG
AGCTGCGTGAAGAACGCCAACCCGCAGTGTGATGGGCGCCCGACTGCAGGGACGGCTCGGA
TGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCTCCAGCCGCATTGTTGGTGGAGCTGTG
CCTCGAGGGTGAGTGCCATGGCAGGCCAGCTCCAGGTTCGGGTGCACACATCTGTGG
GGGGCCCTCATCGCTGACCGCTGGGTGATAACAGCTGCCACTGCTTCCAGGAGGAAGCAT
GGCCTCACGGTGTGTTGGACCGTGTGGCAAGGGTGGCAGAACTCGCGTGGCCTG
GAGAGGTGTCTTCAGGTGAGCCGCTGCTCCCTGCACCCGTACACGAAGAGGACAGCCAT
GAACACGACGTGGCGCTGCTGCAGCTGACCCGGTGGTGCCTGGCCCGTGCAGCCC
CGTCTGCTGCCCGCCGCTCCACTTCTCGAGCCGCCCTGCACTGCTGGATTACGGGCT
GGGGCGCCTTGCCTGGCGAGGGCGGCCATCAGCAACGCTCTGCAGAAAAGTGGATGTGCAGTTG
ATCCCCACAGGACCTGTGCAAGCGAGGCCTATCGCTACCAAGGTGACGCCACGCATGCTGTG
CGGCTACCGCAAGGGCAAGAAGGATGCCGTGCAAGGTGACTCAGGTGGTCCGCTGGTGTG
AGGCACTCAGTGGCCGCTGGTCTGGCGGGGCTGGTCAGCTGGGCTGGCTGGCCGG
CCTAACTACTCGCGCTACACCCGCATCACAGGTGATCAGCTGGATCCAGCAAGTGGT
GACCTGAGGAACGTGCCCCCTGCAAAGCAGGGCCACCTCCTGGACTCAGAGAGCCAGGGC
AACTGCCAACGGAGGGACAAGTAT

FIGURE 65

GGACGAGGGCAGATCTCGTTCTGGGGCAAGCCGTTGACACTCGCTCCCTGCCACCGCCCCGG
CTCCGTGCCCCAAGTTTCATTTCACCTTCTCTGCCCTCAGTCCCCCAGCCCCCTGGCCG
AGAGAAGGGCTTACCGCCGGATTGCTGAAACACCAAGAGGTGGTTTTGTTTTAAA
ACTTCTGTTCTGGAGGGGTGTGGCGGGGCAGGATGACAACTCCGTTCTGCTCTG
TTTCTGGAGCCTCTGCTATTGCTTGCTGCGGGGAGCCCGTACCTTGGTCCAGAGGGAC
GGCTGGAAGATAAGCTCCACAAACCAAAGCTACACAGACTGAGGTCAAACCATCTGTGAGG
TTAACCTCCGCACCTCCAAGGACCCAGAGCATGAAGGATGCTACCTCTCGTCGGCCACAG
CCAGCCCTTAGAAGACTGCAGTTCAACATGACAGCTAAAACCTTTCATCATTACGGAT
GGACGATGAGCGGTATTTGAAAAGCTGGCTGCACAAACTCGTGTCAAGCCCTGCACACAAGA
GAGAAAGACCCAATGTAGTTGTTGACTGGCTCCCCCTGGCCACCAGCTTACACGGGA
TGCGGTCAATAACCAAGGGTGGTGGCACACAGCATTGCCAGGATGCTGACTGGCTGCAGG
AGAAGGACGATTTCTCTCGGAATGTCACCTTGATCGCTACAGCTCGAGCGCACGTG
GCCGGGTATGCAGGCAACTCGTGAAGGGACGGTGGCCGAATCACAGGTTGGATCCTGC
CGGGCCCATGTTGAAGGGGCCGACATCCACAAGAGGTCTCTCCGACGATGCAGATTGG
TGGAATGCTCCACACCTACACGCCTCTCGGCTTGAGCATTGGTATTCAAGATGCCGTG
GCCACATTGACATCTACCCCAATGGGGTACTTCCAGCCAGGCTGGACTCAACGATGT
CTTGGGATCAATTGCAATGGAACAACTCACAGAGGTGGAAAAATGTGAGCATGAGCGAGCCG
TCCACCTTTGTTGACTCTCTGGTGAATCAGGACAAGCCGAGTTTGCTTCCAGTGCACT
GAACCCAATCGCTTCAAAAGGGATCTGCTGAGCTGCCAAGAACCGTTGTAATAGCAT
TGGCTACAATGCCAAGAAAATGAGGAACAAGAGGAACAGCAAATGTACCTAAAACCCGGG
CAGGCATGCCCTTCAGAGGTAACTTCAGTCCCTGGAGTGTCCCTGAGGAAGGCCCTTAATA
CCTCCTCTTAATACCATGCTGCAGAGCAGGGCACATCTAGCCCAGGAGAACGGAAATCTT
CAATCCAATCAAATCGTGCAGGAACTCAGATTACACTGTGCATGTCCTAGGAAAGGGAAATCTT
ACAAAAAAACAGTGTGGACCCCTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

FIGURE 66

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49646
><subunit 1 of 1, 354 aa, 1 stop
><MW: 39362, pI: 8.35, NX(S/T): 2
MSNSVPLLCFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRTSKDPEHE
GCYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVDWL
PLAHQLYTDAVNNTRVVGHSIARMLDWLQEKKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTV
GRITGLDPAGPMFEGADIHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIPNGGDF
QPGCGLNDVLGSIAYGTTIEVVKCEHERAVHLFVDSLNVNQDKPSFAFQCTDSNRFKKGICLS
CRKNRCNSIGYNAAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP
```

Important features:

Signal peptide:

amino acids 1-16

Lipases, serine active site.

amino acids 163-172

N-glycosylation sites.

amino acids 80-83 and 136-139

1011571

FIGURE 67

CGGACCGCTGGCGGACCGCTGGGCCTGGCAAGGGCCGGGCGCCGGGCCAGCCACCTTCTCCCTCCCCGC
TCCCCTGTGCGCCTCGCTGGCTGGAGCGCTGAGGAGTGGAGCACACCCGGCCGGCTGGGGCTGACAGT
CGCGAAAGTTGGCCCGAAGGAGAATGGTCTCAAACCCGGGAGCAGGGCCAGGGCCAGACGGGGCGCTCG
CTGGCTCTGGCGGGCTGTAGGGCGCCGGCTGGAGGACGGGGAGACCCGGGGCTTCAAGAGCCGGCCGGAG
AGAAGAGTGCAGCGGGAGGGAGGGAAACACTCCAAGTGGGAAGGACCGCCCTACTCCGGCTGCG
CGCCGCGCCGGGGCTGGAGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAG
CACCAAGGGAGGCGCTGGCGCCGGGCTCCGGCGACCCCATCGGTAGACAGAGTCGGGACCCCTCG
GCACCTCTGGAGCGCCGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAG
ACCGGATTATTTTCCAAATCATGTGTGGAGGACCCCGAGGGCTCTTAGAAGTGCAGGGACCCCTAAC
GGCCCTGTGGCGGACGGCACCTCCCTGCACACTGCACCTGGCTCATCTGGCAGCAAGGAAGACTG
TCACCATCAGGTTCCAGAACGCTACCTGGCCCTGGCTCAGAGCCTTAACTCAGGCTCCCTCTCAGGAC
TGATCTCCCTGGTGTAGGGAGGACCCCTGGAGCTGGCTGAGCTGGGGGGAGGGAGGGAGGGAG
CTGGGCTGGGAGGACCCATGGGCTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
AGTTTCAGTGTGCTGAAACCCTGGCTGTGTATCTGTGCTGGCTGAGCTGGGTTGATGGCTGTGGCT
CTGATGAAGCAGGGTGTGAGCTCAGACCCCTTCCCTGGCTGACCCCAAGACCCGGTCCCTCCCTG
TCACCTCTGGAGGACTCTATGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
CTGGCCATGGCTGTGGAGCCCCATGATGGCGGGCGCTGGGGCTGCGCTTCAACCCCTGGACTTGG
GAGAGTGCAGTGTGATGTGTGAGGGAGGGCCCTGGGGAGGGAGGGAGGGAGGGAGGGAGGGAG
GCAATTGGCAAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAG
ATGGTCGTTGGCTCAATGCCACCTACATGTGCGGGGATATGCTGCTGCTGGGACAGACCCGTGGCT
CTGGGCTGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAG
GTGCTGACGGCACAGATGGAGGGAGACTGGCCAGCTGGCCAGCTGGCACACTTCCCTGTGGGCT
CTGGCTGGCACACGGGGCTGCTACCTGGCTGTGACCCAGCTGGCACACTTCTCTCTCTCTCT
GAGAGCTGTGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAG
ATGGCAAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAG
CTGGAGTCTGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAG
TTGGCACCACAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAG
CTTCCCTACGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAG
ACTCTAGTGTGGCTACAGGATCTTCACTGGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAG
CCGGCGCTGTGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAG
CCAACA CCCCCTGGGGCTCTGGGGCTGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAG
GTGGCAAGGGTCCAGGGCTGGGGGGGGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAG
AGGGCTCCCTCCCATCTGCTAGACACCTGGCCAGCCGGCCACTCTGGCTGAAGCCGGCCACTGCCCTCAC
TGCCTCTAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGG
CAGGACAAACCGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGG
TGCCACTGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
CTCTACTGGGGCTCTCCCTGGGGCTCTACTCATAGTGTGGCACACACCTTTAGAGGGGGGGGGGGGG
ACCAACTTCTTCCCTGTCTCTGGGGATTTCAAGGGACTTGTGGGGCTCCCGGTGACCCCTATGAGCTG
TAAGTGTGCTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
TCACCAACCCCTGGCTCCCAAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGG
GTTGGAGACTCTCATCTTGGCAAACTCTACCCAAAGTGGGCTTAAAGCACCGGAATGCCAATTAACTAGAGA
CCCTCCAGCCCCAAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGG
CTCACAAAAGAGTGCACAAATGCTTCTATTCCATAGCTACGGCATGGCTCAGTAAGTTGAGGTCAAAAATAA
GGAATCATACATCTC

FIGURE 68

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631
<subunit 1 of 1, 713 aa, 1 stop
<MW: 76193, pI: 5.42, NX(S/T): 4
MLLATLLLLLGGALAHPDRIIFPNHACEDPPPAVLLEVQGTLQRPLVRSRTSPANCTWLIL
GSKEQTVTIRFQKLHLACGGERLTLSPLQPLISLCEAPPSPQLPGGNVTITYSAGARAP
MGQQGFLLSYSQDWLMCLQEEFQCLNHRCVSAVQRCGVDAKGDSDEAGCSSDPFPGLTPRP
VPSLPCNVTLLEDFYGVFSSPGYTHLASVSHPQSCHWLLDPHDGRRLAVRFTALDLGFGDAVH
VYDGPGPPESSRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNNGFNATYHVRGYCLP
WDRPCGLGSLGAGEGLGERCYSEAQRCDGSWDCADGTDEEDCPGCPPGHFPCGAAGTSGAT
ACYLPADRCNYQTFCADGADERRCRHCPGNFRCRDEKCVYETWVCDGQPDCADGSDEWDCS
VVLPRKVITAIVIGSLVCGLLLVIALGCTKLYAIRTQEYSIFAPLSRMEAEIVQQAPPSSY
GQLIAQGAIPPVVEDFTPENPNDNSVLGNLRSLLQILRQDMTPGGPGARRRQRGRLMRRLVR
RLRRWGLLPRNTNPARESEARSQVTPSAPLEALDGTTGPAREGGAVGGQDGQEAPPLPIKA
PLPSASTSPATTVPPEAPGPLPSLPLEPSLLSGVVQALRGRLLPSLGPPGPTRSPPGPHTAV
LALEDEDDVLLVPLAEPGVVVAEAEDEPLLT
```

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

FIGURE 69

CGAGCTGGCGAGAAGTAGGGGAGGGCGGTGCTCCGCCGCGGTGGCGGTTGCTATCGCTTCG
CAGAACCTACTCAGGCAGCAGCTGAGAAGAGTGAGGGAAAGTGTGCTGCTGCTGGGCTGCA
GACGCGATGGATAACGTGAGCGAAAATAAAACATCGCCCTTCTGCTTCAGTGTGAAAGG
CCACGTGAAGATGCTGCGGCTGGCACTA**CTGTGACATCTATGACCTTTTATCATCGCAC**
AAGCCCCCTGAACCATATATTGTTACTGGAAGTCACCGTTATCTTATTTTCATA
CTTTTATATGTA**CTCAGACTTGATCGATTAAATGAAGTGGTTATTTGGCC**TTGCTGATAT
TATCAACTCACTGGTAACACAGTATT**CATGCTCATCGTATCTGTGTTGGCACTGATAC**CAG
AAACCACAA**CATTGACAGTTGGTGGAGGGGTGTTGCACTTGTGACAGCAGTATGCTGTCTT**
GCCGACGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCAGCGGCCTTACCAAAAAAA
GCCTGTGCATGAAAAAAAAGAAGTTTGTAATTTATATTACTTTAGTTGATACTAAGT
ATTAACATATTCTGTATTCTCCAAAAA

FIGURE 70

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645
><subunit 1 of 1, 152 aa, 1 stop
><MW: 17170, pI: 9.62, NX(S/T): 1
MDNVQPKIKHRPFCFSVKGHVKMLRLALTTSMTFFIIAQAPEPYIVITGFEVTVILFFILL
YVLRLDRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTLTVGGGVFALVTAVCCLAD
GALIYRKLLFNPSGPYQKKPVHEKKEVL
```

Important features:

Potential type II transmembrane domain:

amino acids 26-42

Other potential transmembrane domain:

amino acids 44-65, 81-101 and 109-129

Leucine zipper pattern

amino acids 78-99 and 85-106

N-myristoylation site.

amino acids 110-115

Ribonucleotide reductase large subunit protein

amino acids 116-127

101501-101502

FIGURE 71

GGCGGAGAAGTAGGGGAGGGCGTGTCCGCCGCGGTGGCGGTTGCTATCGTTTGACAGAAC
TACTCAGGCAGCCAGNTGAGAAGAGTTGAGGGAAAGTGCCTGCTGGTCTGCAGACGCGA
TGGATAACGTGCAGCCGAAAATAAAACATCGCCCCCTCTGCTTCAGTGTGAAAGGCCACGTG
AAGATGCTGCGGCTGGCACTAAGTGNACATCTATGACCTTTTATNATCGCACAGCCCC
TGAACCATATATTGTTATCACTGGATTGAAGTCACCGTTATCTTATTTTCATACTTTAT
ATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTGGCCTTGCTTGATATTATCAAC
TCACGGTAACAAACAGTATTGTCATGCTATCGTATCTGTGTTGGCACTGATACCAGAAACAC
AACATTGACAGTTGGTGGAGGGGTGTTGCACTTGTGACAGCAGTATGCTGTNTGCCGAC

FIGURE 72

CAGCCCCGCGCGCCGGCGAGTCGCTGAGCCGCGGCTGCCGACGGACGGGACCGGGCTAGG
CTGGCGCGCCCCCGGGCCCGCTGGCATGGCGCACTGGCCCGGCGCTGCTGCTGC
CTCTGCTGGCCAGTGGCTCTGCGCCGCCCCGGAGCTGGCCCCCGGCCCTCACGCTG
CCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCACCCCGGACCCGGACCCC
TGCGAGCGCCACGCCGACGGCTGGCGCTGCCCTGGAGCCTGCCCTGGCGTCCCCCGCG
GCGCCGCCAACCTCTGGCATGGTAGACAACCTGCAGGGGACTCTGGCCGGCTACTAC
CTGGAGATGCTGATCGGGACCCCCCGCAGAAGCTACAGATTCTGTTGACACTGGAAGCAG
TAACCTTGCGCTGGCAGGAACCCCGCACTCCATACAGACAGTACTTGACACAGAGAGGT
CTAGCACATACCGCTCCAAGGGCTTGTAGCTCACAGTAAGTACACACAAGGAAGCTGGACG
GGCTTCGTTGGGAAGACCTCGTACCATCCCCAAAGGCTCAATACTCTTTCTGTCAA
CATTGCCACTATTTTGAAATCAGAGAATTCTTTTGCTGGGATAATGGAAATGAAATAC
TTGGCTAGCTTATGCCACACTTGCCAAGCCATCAAGTCTCTGGAGACCTTCTCGACTCC
CTGGTACACAAAGCAAATCCCCAACGTTCTCCATGCAGATGTGTGGAGCCGGCTTGCC
CGTTGCTGGACTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTGT
ATAAAGGAGACATCTGGTATAACCCCTATTAAAGGAAGAGTGTACTACAGATAGAAATTCTG
AAATTGGAAATTGGAGGCCAACGCTTAATCTGACTGCAGAGAGTATAACGCAGACAAGGC
CATCGTGGACAGTGGCACCCACGCTGCTGCGCCTGCCAGAACGGTGTGATGCGGTGGTGG
AAGCTGTGGCCCGCGCATCTCTGATTCCAAGAATTCTCTGATGGTTCTGGACTGGTCCCAG
CTGGCGTGTGGACGAACTGGAAACACCTGGCTTACTCCCTAAATCTCCATCTACCT
GAGAGACGAGAACTCCAGCAGGTCACTCGTATCACAACTCTGCCTCAGCTTACATTAGC
CCATGATGGGGCCGGCTGAATTATGAATGTTACCGATTGGCATTCCCCATCCACAAAT
GGCCTGGTGTGGTGGACCGGTGATGGAGGGCTCTACGTCATCTCGACAGAGCCAGAA
GAGGGTGGGCTCGAGCGAGCCCTGTGCAAGAAATTGAGGTGCTGCAGTGCTGAAATT
CCGGGCCTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCGCTCAGTCCTTGAGCGAG
CCCATTGGTGGATTGTGTCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCTTGTCTT
AAATCGTCTGCTGCTGCTGCCGTTCCGGTGTAGCGTCGCCCGTGAACCTGAGGTGTC
ATGATGAGTCTCTCTGGTCAAGACATCGCTGAAATGAAAGCCAGGCTGACCTCAAGCAA
CCATGAACCTAGCTATTAAAGAAAATCACATTCCAGGGCAGCAGCCGGATGATGGTGGCG
CTTCTCCCTGTGCCAACCGCTTCAATCTGTTCTGCTCCAGATGCCCTAGATTCAAC
TGTCTTTGATTCTGATTTCAGCTTCAAGCTTCAAATCCTCCCTACTTCCAAGAAAAATAATTAA
AAAAAAAAACTCATTCTAA

FIGURE 73

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493
><subunit 1 of 1, 518 aa, 1 stop
><MW: 56180, pI: 5.08, NX(S/T): 2
MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAATNRVVAPTPGPGTPAERHADGLAL
ALEPALASPAGAANFLAMVDNLQGDSGRGYYLEMLIGTPPKLQILVDTGSSNFAVAGTPHS
YIDTYFDTERSSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATIFESENF
FLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCAGLPVAGSGTNGGS
LVLGGIEPSLYKGDIWYTPIKEEWWYQIEIILKLEIGGGQLNLDCREYNADKAIVDGTLLR
LPQKVFDAVVEAVARASLIPEFSDGFWTGSQQLACWTNSETPWSYFPKISIYLRDENSRSFR
ITILPQLYIQPMMGAGLNYE CYRGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA
EIAGAAVSEISGPFSTEDVASNCVPQAQLSEPILWIVSYALMSVCGAILLVLIVL^{LLL}PFRC
QRPRDPEVVNDESSLVRHRWK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 466-494

N-glycosylation sites.

amino acids 170-173 and 366-369

Leucine zipper pattern.

amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases

amino acids 109-118, 252-261 and 298-310

FIGURE 74

CGCCTCCGCCCTCGGAGGCCTGACCGGCCGGCGCCGTTCCAGGCCGTGCAAGGGCGGATCG
GCAGGCCGCTGGCGCGATCCAGGGCGGTGCGGGGCTGGCGGGAGCCGGAGGCCGCC
GGCATGGAGGCCTGCTGCTGGCGCGGGTGTGCTGGCGCTTA CGTGCTTGTCTACTA
CACACCTGGTGAAGGCCCGCCGTGCGGCCGATGGCAACCTGCGGGGCCGACGCCGTGG
TCACGGGCCAACAGCGGCATCGAAAGATGACGGCCTGGAGCTGGCGCCGGAGCG
CGCGTGGTGTGCTGGCCTGCCGCAGCCAGGAGCGCGGGAGGCGGCTGCCCTCGACCTCGCCA
GGAGAGTGGAAACATGAGGTATCTTATGCCCTGGACTTGGCAGTCTGCCCTCGGTG
GGGCCTTGGCACTGCCCTCTGAGCTCTGAGCCACGGTGGACATCCTCATCCACAATGCC
GGTATCAGTCTCTGTGGCGGACCGTGAGGCCTTAACCTGCTGCTCGGGTGAAACCATA
CGGCCCTTCTGCTGACACATCTGCTGCTGCCCTGCCGAAAGGCATGTGCCCTAGCCGCG
TGGTGGTGTAGCCTCAGCTGCCACTGTGGGAGCTTGTGACTTCAAAGCCTGGACCGC
CCAGTGGTGGGCTGGCGCAGGAGCTGCGGGCATATGCTGACACTAAGCTGGTAATGTACT
GTTTGCCGGGAGCTGCCAACAGCTTGAGGCCACTGGCGTCACCTGCTATGCAGCCCACC
CAGGGCCTGTGAACTCGGAGCTGTTCTGCGCCATGTTCTGGATGGCTGCGCCACTTTG
CGGCCATTGGCTGGCTGGCTGGCTCGGCCAACAGGGGGTGGCCAGACACCCCTGTATTG
TGCTCTACAAGAGGGCATCGAGGCCCTCAGTGGAGATATTTGCCAACTGCCATGTGGAG
AGGTGCCCTCAGCTGCCAGACGCCAGGGCAGCCATGCCCTATGGAGGCGAACAGAGG
CTGGCAGGGCTGGGCTGGGAGGATGCTGAACCCGATGAAGACCCCAGTCTGAGGACTC
AGAGGCCCATCTCTCTAACGCCACCCCCCACCTGAGGAGGCCACAGTTCTCAACCTTACC
CCAGGCCCTCAGAGCTCACCAAGATTGCTAAAGATGACGCCAGGAATTCCAGGCTAAAGTTGAG
CCTGAGATCCAGCTCCTAACCCTCAGGCCAGGATGCTGCCATGCCACTTCTGTTGCC
GAAAACCTGGATGTGTGAGGCCATGCCCTGGACACTGACGGGTTGTGATCTTGACCTC
CGTGGTTACTTCTGGGCCCCAACGCTGTGCCCTGGACATCTCTTCTGGTTGAAGGAAT
AATGGGTGATTATTTCTCTGAGAGTGACAGTAACCCAGATGGAGAGATAGGGTATGCT
AGACACTGTGCTCTCGGAAATTGGATGTAGTATTTCAGGCCACCCCTTATTGATTCTG
ATCAGCTCTGGAGCAGAGGCAGGGAGTTGCAATGTGATGCACTGCCAACATTGAGAATTAG
TGAACCTGATCCCTTGCAACCGCTAGCTAGTAGGTAGTTAAATTACCCAGTGTAAATGAAGCG
GAATTAGGCTCCCGAGCTAAGGGACTGCCCTAGGGTCTCACAGTGAAGTAGGAGGAGGGCCTG
GGATCTGAACCCAAGGGCTGAGGCCAGGGCAGACTGCCGAAGATGGGTGCTGAGAAGTGA
GTCAGGGCAGGGCAGCTGGTATCGAGGTGCCCATGGAGTAAGGGACGCCCTCCGGCG
ATGCAGGGCTGGGTATCTGTATCTGAAGCCCTCGGAATAAGCGCGTTGACGCCAAAA
AAAAAAAAAAAAAAA

FIGURE 75

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48227
<subunit 1 of 1, 377 aa, 1 stop
<MW: 40849, pI: 7.98, NX(S/T): 0
MEALLLGAGLLLGAYVLVYYNLVKAPPCCGMGNLRGRTAVVTGANSIGKMTALELARRGAR
VVLACRSQERGEAAAFDLRQESGNNEVIFMALDLASLASVRAFATAFLSSEPRLDILIHNAG
ISSCGRTREAFNLLLRLRVNHIGPFLLTHLLLPCLKACAPSRRVVVASAAHCRGRLLFKRLDRP
VVGWRQELRAYADTKLANVLFARELANQLEATGVTCYAAHPGPVNSELFLRHVPGWLRPLL
PLAWLVLRAPRGGAQTPLYCALQEGIEPLSGRYFANCHVEVPPAARDRAHRLWEASKRL
AGLGPGEDAEPDEDPQSEDSEAPSSLSTPHPEEPTVSQPYPSPQSSPDLSKMTHRQAKVEP
EIQLS
```

Important features:

Signal peptide:

amino acids 1-16

Glycosaminoglycan attachment site.

amino acids 46-49

Short-chain alcohol dehydrogenase family

amino acids 37-49 and 114-124

0001785444.101601

FIGURE 76

FIGURE 77

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404
<subunit 1 of 1, 985 aa, 1 stop
<MW: 105336, pI: 6.55, NX(S/T): 7
MGGMQAQDSSPPQILVHPQDQLFQGPGPARMSCQASGQPPTIRWLLNGQPLSMVPDPHHLPP
DGTLLLQPPARGHAHDGQALSTDLGVTCEASNRGLGTAWSRGARLSVAVLREDFQIQPRDM
VAVVGEGQFTLECGPPWGHPEPTVSWWKDGPALQPGRHTVSGGSLLMARAEEKSDEGTYMCV
ATNSAGHRESRAARVS1QEPEQDYTEPVELLAVRIQLENVTLLNPDAEGPKPRPAVWLWKV
SGPAAAPAQSYTALFRQTAPGGQQGAPWAEELLAGWQSAELGGLHWGQDYEFKVRPSSGRARG
PDSNVLLRLPEKVPSPAPPQEVTLKPGNGTVFVSWVPPAENHNGIIRGYQVWSLGNTSLPP
ANWTVVGEQTQLEIATHMPGSYCVQVAVTGAGAGEPSRPVCCLLEQAMERATQEPESEHGPW
TLEQLRATLKRPEVIATCGVALWLLLGTAVC1HRRRARVHLGPGLYRYTSEDAILKHRMD
HSDSQWLADTWRSTSGSRDLSSSSSLSSRLGADARDPLDCRSLLSWSRSPGVPLLPDTST
FYGSLIAELPSSTPARPSPQPVAVRRLPPQLAQLSSPCSSSDLCSRGLSSPRLSLAPAEEA
WKAKKKQELQHANSSPLLRGSHSLELRACELGNGRSKNLSQLSPGAVPQALVAWRALGPKLLS
SSNELVTRHLPPAPLFPHEPTPPTQSQQTQPPVAPQAPSSILLPAAPIPILSPCSPPSPQASS
LSGPSPASSRLSSSSLSSLGEDQDSVLTPPEEVALCLELSEGEETPRNSVSPMPRAPSPTTY
GYISVPTASEFTDMGRTGGVGPKGGVLLCPPRCLTTPSEGSLANGWGSASEDNAASARA
SLVSSSDGSPFLADAHFARALAVAVDSFGFGLPREADCFVIDASSPPSPRDEIFLTPNLSLP
LWEWRPDWLEDMEVSHTQRLGRGMPPWPDSQISSQRSQLHCRMPKAGASPVDYS
```

Important features:

Transmembrane domain:

amino acids 448-467

N-glycosylation sites:

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

N-myristylation sites.

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

Phosphotyrosine interaction domain proteins

amino acids 740-753

FIGURE 78

CTCCCACGGTGTCCAGCGCCCAGA**AT**GGGCTTCTGGCTATGGGTTGCCCTGCTGCT
CCCAGGTTATGAAGCCCTGGAGGGGCCCAGAGGAATCAGGGGTTCGAAGGGGACACTGTG
CCCTGCACTGACCTACAGGGAAAGAGCTGAGGGACCACCGAAGTACTGGTCAGGAAGGGT
GGGATCCTCTCTCGCTCTGGCACCATCTATGAGAAGAAGGCCAGGAAGAACAT
GAAGGGCAGGGTGTCCATCGTGACAGGCCAGGAGCTCTCGCTCATGGTGAACCTGTGGA
ACCTCACCCCTGCAAGACGCTGGGGACTACTGGTGTGGGTCGAAAAAGGGGCCCGATGAG
TCTTTACTGATCTCTCTGCTTCTCAGGACCCCTGTCCTCCCTCCCTTCTCCAC
CTTCCAGGCCCTGGCTACAACAGGCCCTGCAGCCCAGGAAAAGCTCAGCAAACCCAGCCCC
CAGGATTGACTCTCTGGCTTACCCGGCAGGCCACACGGCCAAGCAGGGAAAGACAGGG
GCTGAGGCCCTCCATTGCCAGGGACTTCCAGTACGGGACGAAAAGACTTCTCAGTACAC
AGGAACCTCTCCACCCAGCAGCCTCTCCTCTGCAAGGAGCTCCGCCCCCCCATGAGC
TGGAAGCTCCACCTCAGCAGGGACACCGACTCAGCAGTGGCAGCTAACGGCCAG
GTGTCATCCCCATGGTCCGATCATGGGCCCCAGTCTCTGGTGTGCTGAGCCCTCTGTCAGC
CGCAGGCCCTGATCGCCTCTGCAAGCCACTGCTCTGTGGAGAAAGGAAGGCTCAACAGGCCA
CGGAGAACAGAGGAAACAGAGAAGTCTGGCTCTACGGCTTACTGCCAGGGAAAGGAAGGCC
CTTCCAGGCCCTGGGAGCTGATCTGATGCCCTCCACACATCTGAGGAGGA
GCTGGGCTTCTGAGTTGCTCTAGCTGAGGGCAGGCCCTCTGCCAGGCCAGCAG
GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAAGCTTCCACCTCAGCTCAGAG
TCCAGCTGGCCGACTCAAGGGCTCTCCCAACCTCCCTCCAGGCTCTCTTGCATGTTCCA
GCCGTACCTAGAGCTGGTTCTGAGCCCTGGAGGCCAGAGCGGTGGCCTTGTCTTCCGGCTG
GAGACTGGGACATCCCTGATAGGTTCACATCCCTGGCAGAGTACCAAGCTGCTGACCCCTCA
GCAGGCCAGAACAGGCTAGGGATCTGGCTGAGTTCAATCTGCCAGGAACCTCTGGC
CTCTAGGCCCTGCGGCCCTCTCCACTCTCAGCACGGGCCCCACCTTGTCTTCCCTCC
TGGGCTCTGAGACTTACTGCTCCACGGCTCTGAGCTCATCGTGGTGTGATGAAGAGGACATGCT
GGGGTGGAGACTGGGATTCTGGCTCTCTTGAACCACCTGCACTCCAGGCCCTCAGGAAGGCT
GTGAAAAACCTGAGACTCTGGCCCTCCACCAAGGCCACCAAAACCATCTGGGCTTGGCTGAG
GACTCTGAAATCTAACATGGCCCTGACTGCTGCACTTGTGAGTTGAGGGCCAGGCCCTG
ATGAACGTCACACCCCTCAGCTTAGAGTCTGCATTGGCTGTGAGCTCTCCACCTGCC
CAATAGATCTGCTCTGCTGCAACAGATCCACGTTGGGACTCCCTGAGGCTGCTAAG
TCCAGGCCCTGGCTGGCAGGTGACCATGAGATAAGGCCAGGACGGGCCACAGAAGTGG
TTGCCCTTNCCATTGCCCTCTGGNCATGCCCTTGTGCTTGGAAAAATGATGAAGA
AAACCTTGGCTCTCCCTGTCTGAAAGGGTTACTTGCCTATGGGTTCTGGCTAGAGA
GAAAAGTAGAAAACAGAGTGCAGCTGGTAGTGTCAACACAGAGGAGAGTAGGAACAGGGCG
ATACCTGAAGGAGTCACTGGCAGTCCAGGCCCTGGAGAAGGGCTGGGTTGGTAAAGTA
GCACAACACTATTTTTCTTTTCCATTATTATTGTGTTTTAAAGACAGAACTCTGTGCT
GCTGCCAGGCTGGAGTGGCAGTGGCACATCTGAAACCTCCGCCCTCTGGGTTCAAGTGGATT
CTTCTGCCCTAGCCCTCCAGTAGCTGGGATTACAGGCCAGCACCCACACCTGGCTAATT
TTTGACTTTAGTAGAGATGGGTTTACCATGTTGCCAGGCTGGCTTGAACCTCTGAC
CTCAAATGAGCCCTGCTTCAGTCTCCAAATGCCGGGATTACAGGAGTGGCAACTGTG
TCTGGCCCTATTCTTTAAAAAGTGAATTAAGAGTTGTTCACTGAGTATGCAAAACTTGGAAAG
ATGGAGGAGAAAAGAGAAAAGAGTCACTGGCCCTTGTGCTACCCAGAGACTATCAT
TATTTCTGTTTGTGACTCTCTTCCACTCTTTCTTCTTCAACATAATTGCCGGTGTCTT
TTTACAGAGCAATTATCTGTATATAACACTTTGTATCCCTGCCCTTACACCTTATGTTCC
ATCACTTATTCCAGCACTCTGTTTACAGACCTTTATAAATAAAATGTTCATCA
GCTGCATAAAAAAAAAAAAAA

FIGURE 79

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196
<subunit 1 of 1, 332 aa, 1 stop
<MW: 36143, pI: 5.89, NX(S/T): 1
MRLLVLLWGCLLPGYEALEGPEEISGFEGDTVSLQCTYREELRDHKYWCRKGGLLFSRCS
GTIYAAEEEGQETMKGRVSIRDSRQEELSIVTLWNLTQDAGEYWCVEKRGPDSSLISLFV
FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTPGLYPAATTAKQGKTGAEAAPPLPG
TSQYGHERTSQYTGTSPHPATSPPAGSSRPPMQLDSTSaedtspalssgsskprvsiPMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETORNEKFWLRLTAEEKEAPSQAPEGD
VISMPPLHTSEEELGFSKFVSA
```

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

FIGURE 80

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCCGTGGGGTGGCAGGA
GCCGCAGGCCAGAGCAGACAGCCGAGAACAGGTGAGCTGCAGCTCAGACTCT
GCTCTGTTGCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCT
GGACTTGAGAAATCCTCCTGCCTTAGCCTCTGCATATCTGGGACTCCAGGGTGCACCAA
GCCCTGTTCTCTCCTCTGTGAGTGGACCACGGAGGTGGTGGCTCACCTCTGCCGCCGCATAGAACCCAA
AGCTCAGCTCTGAGCCAGAGTGGTGGCTCCACCTCTGCCGCCGCATAGAACCCAA
CAGGGCTCTCAGAAGGCCGGTGGTGGCCAGCTGGGATCATGTTGTTGGCCCTGGTCTGTCTGC
TCAGCTGCCCTGCTACCCCTCAGTGAGGCCAAGCTCAGGTGTTGTACTGGCCAGAGTG
CTACATGACTTCGGCTGGACGGATAACGGGATAACGCCCTGGCTGACTGGGCTGCCCTGC
TTATTTCAACAAGCGTTTCAACGCAGCTCTGGACTACGAGGCTGATGGGAGCACCAACA
ACGGGATCTCCAGATCAAACGCCGGAGGTGGTGCAGCAACCTCACCCGAACGTCCCCAAC
GTGTGCCGGATGTACTGCTCAGATTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGC
CATGAAGATAACCCAAGAGGCCCTCAGGGCTGGGTTACTGGGAGGCCTGGAGGCATCACTGCC
AGGGAAAAGACCTCACTGAATGGGTGGATGGCTGTGACTTCAGGATGGACGGAAACCATGCA
CAGCAGGCTGGAAATGTGGTTGGTCCGTGACCTAGGCTTGGGAAGACAAGCCAGCGAATA
AAGGATGGTGAACGTGAAA

FIGURE 81

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187
<subunit 1 of 1, 146 aa, 1 stop
<MW: 16430, pI: 5.05, NX(S/T): 1
MLLALVCLLSCLLPSSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALD
YEADGSTNNNGIFQINSRRWCNSNLTPNVPNVCRMYCSDLLNPNLKDTVICAMKITQEPMQGLGY
WEAWRHHCQGKDLTEWVDGCDF
```

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homologous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

FIGURE 82

AGCCGCTGCCCGGGCGCCCGGGCAGGGCACCTGAGTCGGCTCGTGCCTGCGTTC
GCTGCGCCTCTCGTCCTCGCCGTCTCTCAGCGCCGAGCACTGGCTGTACCTGGCCA
AGCTGTCGTCGGTGGGAGCATCTCAGAGGAGAGCTGCGAGAAACTCAAGGGCTGATC
CAGAGGCAGGTGCAAGCGAACCTGGAAGTCACTGGACTCGGACTGGCGCTGGAACTGCTCCACAC
CCAGCTGGCATTGAGGAGTGCCAGTACACAGTTCCGAAACCGGCGCTGGAACGGACTCGGAGGGCGGCTTCGTG
TCGACTCCTGCCCGTCTCGCAAGGGTGTGACGCAAGGGACTCGGAGGGCGGCTTCGTG
TAGCCCATCTTCGGCAAGGTGTGGCCTTGCACTGACGCGGGCGTGCAGCAGTGGGAGCT
GGAGAAGTGGGCTGTGACAGGACAGTGCATGGGTCAAGCCCACAGGGCTTCAGTGGTCAG
GATGCTCTGACAACATCGCTACGGTGGCCTCTCACAGTCGTTGTGGATGCGGGAG
AGAAGCAAGGGGCCCTGTCAGCAGAGCCCTCATGAACCTCCACAACAATGAGGCCGGAG
GAAGGCCATCTGACACACATGCCGGTGAATGCAAGTGCACGGGACGGGTGTCAGGCTCTGTG
AGGTTAAAGACGTGCTGGCGAGCGTGCCGGCCCTCCGCCAGGTGGGTACGCACAGAAGGAG
AAGTTTGATGGTGCACACTGAGGTGGAGGCCACGCCGCGTGGCTCCTCCAGGGACTGGTACCC
ACGCAACGCACAGTCAACCGCACACAGATGAGGACCTGGTGTACTGGAGCCTAGCCCCG
ACTTCTGTGACAGGACATGCGCAGCGCGTGTGGGACAGGGGCCACATGCAACAAG
ACGTCAGGCCATCGAGGCTGTGAGCTGCTGTGCTGCGGCTTCCACACGGCGCA
GGTGGAGCTGGTGAACGCTGCAAGCTGCAAAATTCCACTGGTGTGCTCGTCAAGTGCAGG
AGTGCACAGGGCTCGTGGAGTTGACACGTGCCGATTGACCGGCTGCCTAGCCCTGCCGGC
AAACACCTAGTGGCCAGGGAGGCCATAATTAAACAGTCTCCCACACCTACCCAAAGA
GATACTGGTTGTATTTTTGTCTGGTTTGGTTTGGGTCTCATGTTATTTATGGCGAA
ACCAGGCAGCAACCCCAAGGGCACCAACCAAGGGCTCCCCAAAGCTGGGCTTGTGGCT
GCCACTGACCAAGGGACCTTGCTCGTGCCTGGCTGCCGATGTGGCTGCCACTGACCA
CTCAGTTGTATCTGTGTCGTTTCTACTTGCAAGACCTAAGGTGGAGTAACAAGGAGTAT
TACCAACACATGGCTACTGACCGTGTCACTGGGAAGAGGGGCTTATGGCAGGGAAAATA
GGTACCGACTGTGATGGAAGTCACACCCCTGGAAGGGAAAAGAAACTCTTAACCTCCAGCACACA
TACACATGGACTCTGGCAGCTTGAGGCTAGAACGCCATGTCCTCAATGCCCTGAGAAAGG
GAACAAGCAGATAACCAAGGTCAAGGGCACCAAGGGTCACTTCAGCCCTACATGGACAGCTAGA
GGTTGATATCTGTGGGCTTCCAGGCAAGAAGAGGGAGATGAGAGCAAGAGACGACTGAA
GTCCCACCCCTAGAACCCAGCCTGCCCTGGCAAGGAAACTTAACCACTCC
CCAGACCCACCTAGGCAGGCATATAGGTGCCATCTGGACCAGGGATCCGGCTGTGCTT
TGCAGTCATGCCAGTCACCTTCAACAGCCTGTTCTCCATGAAACTGAAAAACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
GAGAGGGAGGAAAGGGCTGTGCTTGCAGTCATGCCAGTCACCTTCAACAGCACTGTTCTC

FIGURE 83

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48328
<subunit 1 of 1, 351 aa, 1 stop
<MW: 39052, pI: 8.97, NX(S/T): 2
MSPRSLRSLRLVFAVFSAASNWLYLAKLSSVGSISEEETCEKLKGLIQRQVQMCKRNLE
VMDSVRRGAQLAIEECQYQFRNRRWCSTLDSPVFGKVVTQGTREAAFVYAISSAGVAFAV
TRACSSGELEKCGCDRTVHGVSPOGFQWSGCSNDIAYGVAFSQSFDVRERSKGASSRALM
NLHNNEAGRKAITLEMRVECKCHGVSGSCEVKTCWRAPPPFRQVGHALKEKFDGATEVEPRR
VGSSRALVPRNAQFKPHTDEDLVYEPSPDFCEQDMRSVGLGTRGRTCNKTSKAIDGCELLC
CGRGFHTAQVELAERCSCKFHCCFKCRQCQRLVELHTCR
```

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 88-91 and 297-300

Wnt-1 family signature.

amino acids 206-215

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

FIGURE 84

CGGACGCGTGGCGGACGCGTGGCGGACGCGTGGCGGACGCGTGGGTGCCTGCAT
CGCCATGGACACCACCAAGGTACAGCAAGTGGGCGGCAGCTCCGAGGAGGTCCCCGAGGGC
CCTGGGGACGCTGGGTGACTGGAGCAGGAGACCCCTCTTCTGGCCCTGGCTGTCTGGTC
ACCACAGTCCCTTGGGCTGTGATTCTGAGTATCTATTGCTCAAGGCCACGGAGCGCGC
GGCGCTGCTTGACGGCACGACCTGCTGAGGACAAACCCCTCGAAGCAGACGGCGGGCTGG
GTGCCCTGAAGGAGGAGGTGGAGACTGCCACAGCTGCTGCCGGGACGCAGGCGCAGCTG
CAGACCAACGCGCGGGAGCTTGGGAGGGCGAGCGAAGCTGATGGAGCAGGAGAGCGCCCT
GCGGAACTGCGTGAGCGCGTGAACCAAGGGCTTGGCTGAAGCGGGAGGGCGTGGAGGAC
TCCGCACTGAGCTGTTGGCGCTGGAGGCGTGAGGCTCCAGAACAACTCTGCGAGCCG
TGCCCCACGTCGTGGCTGCTTCGAGGGCTCTGCTACTTTCTGTGCCAAAGACGAC
GTGGGCGGGCGGCAGGATCACTGCGCAAGTGCACGGCCACCTGGTGATCGTGGGGCC
TGGATGAGCAGGGCTTCTCACTCGGAACACGCGTGGCGTGGTACTGGCTGGGCTGAGG
GCTGTGCGCCATCTGGCAAGGTTCAAGGCTACCGTGGCTGGGAGGCTCTCTCAGCTT
CAGCCACTGGAACCAGGGAGAGGCCAATGACGCTTGGGCGCGAGAACACTGTGTCATGATGC
TGCACACGGGCTGTGGAACGACGCCAGGTGACAGCGAGAAGGACGGCTGGATCTGTGAG
AAAAGGCACAACTGCTGACCCCGCCAGTGCCTGGAGCCGCAGCCATTGCAAGCATGTCGTA
TCCTGGGGCTGCTCACCTCCCTGGCTCTGGAGCTGATTGCCAAAGAGTTTTCTTCCT
CATCCACCGCTGCTGAGTCTCAGAACACTTGGCCAACATAGCCCTGTCAGGCCAGTGC
TGGGCTCTGGACCTCCATGCCACCTCATCCTAACTCCACTCACGCAGACCCAACCTAAC
TCCACTAGCTCAAATCCCTGCTCTGGCTCCCGTGATATGCCCTCACTTCTCCCTAA
CCAAGGTTAGGTGACTGAGGACTGGAGCTGGAGCTTTGGTTCTCGCATTTCCACCAA
AGCTGTTTGCAGCTGAGGAAGCATATAAAATTGAGAAATGAAAAAA

FIGURE 85

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56352
<subunit 1 of 1, 293 aa, 1 stop
<MW: 32562, pI: 6.53, NX(S/T): 2
MDTTRYSKWGGSEEVPGGPWGRWVHWSSRPLFLALAVLVTTVLWAVILSILLSKASTERAA
LLDGHDLLRTNASKQTAALGALKKEVGDCHSCCSGTQAQLQTTRAELGEAQAKLMEQESALR
ELRERVTQGLAEAGRGRGREDVRTELFRALEAVRLQNNSCEPCPPTSWLSFEGSCYFFSVPKTTW
AAAQDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLFS
HWNQGEPNDAWGRENVCVMLHTGLWNDAPCDSEKDGVICEKRHNC
```

Important features:

Type II transmembrane domain:

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

FIGURE 86

GCCAGGGGAAAGAGGGTATCCGACCCGGGAAGGTCTGGCAGGGCAGTTGGAAAGCG
GCAGCCCCCGCCGCCGGGGAGCCCCTTCCTCCCTTCTCCACGTCCTATCTGCCCTCG
CTGGAGGCCAGGCCGTGCAGCATCGAAGACAGGGAGGAAGCTGGAGCCTATTGGCCCGGG
GGCGCCGGCTCGGGCTTAAATAGGAGCTCGGGCTCTGGCTGGGACCCGACCGCTGCCGG
CGCGCTCCCGCTGCTCTGGCGGGTGATTGAAAACCCCAGCCCGCGCCCTGGGCAAG
GCCCTCTGCGCTCTCTCTGGGCACTCTCGGCGCCGGCCAGCCTTGGGGAGAGTC
CATCTGTTCCGCCAGGCCCCGGCAAATACAGCATTACCTTCACGGCAAGTGGAGCCAGA
CGGCCCTTCCCAAGCAGTACCCCCCTGGCTCGGCCCCCTGGCAGTGCTTCGCTGGGG
GCCGCGCATAGCTCGACTACAGCATTGAGGAAGAACCGTAGCTCAGTAACGGCTGCC
CGACTTTGGGAGCGGGCGAGGGCTGGGCTGTAGAAGGAGATCGAGGGCGGGGGAGG
CGCTGCAAGCGTGCACGAGGTGTTTCGGGCCCGCTCCCCAGCGGACCGGGAGACAG
TCGGCGAGCTGGAGGTGCAAGCGCAGGACTCGCTGGCTCGTTGTGGTGGCAGTCG
CAGCCCCGACTGGTCTGGGGCTGGACAGCCTGGACCTGTGCGACGGGACCGTTGGGG
AACAGGGGGCGTGGACCTGTACCCCTACGACGCCGGAGCGACAGGGCTCACCTCTCC
TCCCCCAACTTCGCCACCATCCCGCAGGACACGGTGACCGAGATAACGTCCTCTCTCCAG
CCACCCGCCAACTCTCTACTACCCGGCTGAAGGCCCTGCCCTCCATGCCAGGGTGA
CACTGCTGCCGCTGCGACAGAGCCCCAGGGCTTCATCCCTCCGCCAGTCTGCCAGC
AGGGACAATGAGATTGTAGACAGCGCCTCAGTTCCAGAAACGCCGCTGGACTGCGAGGTCTC
CTGTGGTGTCTGGGACTGTGCGGAGGCCACTGTGGAGGTCTGGGACCAAGAGCAGGA
CTCGCTACGCTCCGGTCCAGCCGCAACACGGGAGCCCTGCCAGTCGAAGAACAG
GCTGAGTGCCTGATTAACGCGCTTAAAGCCAGAGCCCCGAGCCCTGGGGCCCCCG
GAGCCATGGGTGTCGGGGCTCTGTGCAAGGCTCATGCTGCAGGCCGAGGGCACAGGG
GGTTTCGCCGCTCTGACCCGGTGAAGGCCGCCACATCTGCACTGAAGGGCCCT
CTGGTGGCCGGCACGGCATTGGAAACAGCCTCTCTCTTCCACCTGCTTCTTGGGG
CCCCCGTGTCCCGTCTGCTCTCGCCTCTCCCTGCAAGGATAAGTCATCCCAAGGCTC
CAGCTACTCTAAATTATGTCCTTATAAGTTATTGCTGCCAGGAGATTGCTTCATCG
TCCAGGGGCTGGCTCCACGTGGTGCAGATAACCTCAGACCTGGTCTCTAGGCTGTGCTG
AGCCCACCTCCCGAGGGCGATCCAAGGGGGCCACTGAGAAGTGAATAAAATGGGGCG
TTTGGAAAGCGTCACTGGTTCATGTTATGGATCTCTGCTGGTTGAATAAAAGACTATCTCT
GTTGCTCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 87

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53971
><subunit 1 of 1, 331 aa, 1 stop
><MW: 35844, pI: 5.45, NX(S/T): 2
MENPSPAAALGKALCALLLATLGAAGQPLGGESICSARAPAKYSITFTGKWSQTAFPQYPL
FRPPAQWSSLGAAHSSDYSMWRKNQYVNSNGLRDFAEERGEAWALMKEIEAAGEALQSVHEVF
SAPAVPSGTGQTSAELEVQRRHSLVSFVVRIVPSPDWFVGVDSDLCDGDRWREQAALDLYP
YDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSHPANSFYYPRLKALPPIARVTLRLRQSP
RAFIAPPVLPSPRDNEIVDSASVPETPLDCEVSLWSSWGLCGGHCGRLGTSRTRYVRVQPA
NNGSPCPELEEEAECVPDNCV
```

Important features:

Signal peptide:

amino acids 1-26

FIGURE 88

GGCGGCGTCGTGAGGGCTCCCTTGGGCAGGGTAGTTGGTGCCTGTCTGCGTA
TATTGACAAACTGAAGCTTCCTGCACCCTGGACTTAAGGAAGAGTGACTCGTAGGCCA
CAGCTTCTAGGGCGCCGCCGCTCATCCCCGTAAGGAGCAGACTCTTGTACTGAC
CAAGATGAGCAACATCTACATCCAGGAGCCTCCACGAATGGAGGTTTATTGAAA
CAGCTGGAGATATTGACATAGAGTTGTTCCAAGAAGCTCTAACGCTTGAGA
ATCCAACCTTGTGGAAAGCTTATTGACAATACCATTTTCTAGAGTTGTCCTGGTTT
CATAGTCCAAGCGGAGATCTACTGGCACAGGGAGTTGGAGAGTCTATCTATGGAGC
CATTCAAAGATGAATTCTACCGGTTGCTTAACTGGAGAGGACTGGTGCATGGCA
AATGCTGGITCTCATGATAATGGCAGCAGTTTCTCACACTGGTCAGCAGATGAACT
TAACAATAAGCATAACCATTTGGAAAGGTTACAGGGATACTAGTATATAACATGTC
TGTCAAGAGTAGACATTGATGAGCAGAACACATAATCCACACAAAAAAAGCTGT
GAGGTTTGTAACTCCTTGATGACATCATTCAAGGGAAATTAAAAGGCTGAAAAAAGA
GAAACCCAGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAAATTAGTTACTTT
CATTGGAGAGGAAGCTGAGGAAGAAGAGGGAGGAGTAATCGAGTTAGTCAGAGCATGA
GGCAAAAGCAAAGTAGTCATGACTTGTCAAGGATGATCCACATCTCAGTTCTGTC
TGTAGAAAGTGAAGAAAGGTGATGCACCAGATTAGTTGATGATGGAGAAGATGAAAGTGC
AGCATGATGAATATATTGATGGTGTGAAAGAACCTGATGAGAGAAGAATTGCA
TTAAAAAAGACACAAGTGCAGTAAATCAGCTGGAGAAGGGAGTAACTGAGTTAGTC
AGTCAACAGCTAAAGAAGGGAACTTCCGGGAAGATCAGACCCCTGCACTGCTGA
TTAACATCTAAACTCACTCAAGCAATTGTCGAAACACCTGAAAATGACATTCTG
GTAGAAGATGATGAAGGATGGATGTACATGACTTCAGTTGAGGATAAAAGCAGAAAGT
GAAAGATGCAAGCATGCAAGACTCAGATACATTGAAATCTGATCCTCGGA
ATAAAAGAAGGAGGGAAAGAACAGCAAAAGCTGATGAGAGAGAAAAAGAAAAGAAG
GAGAATAATGATAACCGAACTTGCTGGAAATGTGCCTACAATGGCCTGTAACAGC
TTCCCAACAGCATCACTTAGGGGTGTGAAAAGAAGTATTGGAAACCTGTTGCTGG
AAAAACAAATTATCTTGTTGCAAATTGGAATGATGTAAGCAAATGCTTGGTTACTG
TACATGTGTTTCTAGCTGACCTTTATATTGCTAAATCTGAAATAACTTCC
TCCACAAAAAA

FIGURE 89

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50919
><subunit 1 of 1, 472 aa, 1 stop
><MW: 53847, pI: 5.75, NX(S/T): 2
MSNIYIQEPPNGKVLKTAGDIDIELWSKEAPKACRNFSQLCLEAYYDNTIFHRVVPGFIVQGGDPTGTGSGGESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFTLGRADELN
NKHTIFGKVTGDTVYNMLRLSEVDIDDDERPHNPNPKIKSCEVLFNPFDIIPREIKRLKKEK
PEEEVVKLKPKGTKNFSLLSFGEAAEEEEEVNRVSQSMKGKSCKSHDLLKDDPHLSSVPVV
ESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKKLKDKTSANVKSAGEGEVEKKSV
SRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEEAPPDGAVAELYRREKQKYEALRK
QQSKKGTREDQTLALLNQFKSKLTQAIATPENDIPETEVEDDEGWM SHVLQFEDKSRKVKDASM
DASMQDSDTFFIYDPRNPVNKRREESKKLMREKKERR
```

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 109-112 and 201-204

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.

amino acids 49-66

Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase

amino acids 96-140, 49-89 and 22-51

FIGURE 90

CGCCGCGCGTGGGGCTGGAAAGTCCCGCAGGTCCGTGCCGGCAGAGAGAGATGCTGCCGG
CCCGCCTGGCTTGGAGGGCAGAGAAGTGTCCCAGACCACTTTCGCTTGACTGCCGCTCG
AGCCCTGGCAGA~~AT~~TGTCACAGGGTTCTCTTCTCGGGACTCTGGGCTCCAACACC
TGCGGCCGGCGGGGACGACACAGGGCGGTTTCTCTTCCGGAACGGGAACGTCAGCAA
CCCTCTGGGGCTCAATTGGAAATCTGGAAAGTACTTCAACTCCAGCAACTACATCTG
CTCTTCAAGTGGTTTGAAACGGGGCTTCTGGATCTAAACCTGCACCTGGGTAACCTCTA
GGAGGAACAAATAACAGTGGCTTGCACACAAGGGCTCAAGTGGTACCAAATAATGGAA
CCTGCAAGGAAAACAGATGATGTGGGGAGACACCCATCCAAGTCTTTAGGAGTCCCCCT
TCTCCAGACTCTCTAGGTTACCTCAAGGTTTGACCTCCAGAACCCCCGGAGGCCCTGGAAA
GGAATCAGAGATGATGTACCAACCTACCCGGCTGATGGAGTCCTGCTCTGAGCTGGAG
TCAGTGGCACGATCTGGGCTCACTGCACACCTCCGGCTTCAAGCGAGTCTCTGCG
CTCAGCCTCTGAGTGTCTGGGCTACAGGTGCCGTGAGGAGTCTGGGCCAGCTGGCTCG
ATGTAAGTCTCAGCACCCGGAAACGGTAACTGGGCTGCTGGCTTCAAGCGAGGACTGTCTG
GAACGTTGACCGCGCCGGCGCCGGGGATCCCCAGTGGCTGAGTGGTCTGGTTC
CGGGAGGCCCTTCTATCGTGGGCGCTCTCTCTGACAGGGCTCTGACTTGGGGCCCG
GAGAAACTGGTGTGGTTCTGAGCACAGGCTGCGCATCTGGCTTCTGAGCACGGGA
CGACAGCCACCGCGCCGGAAACTGGGGCTGGGAGCATGGGCTCTGGCTGGTGC
AGGAGAAACATCGCAGCTCTGGGGAGACCAAGGAAATCTGACCTGGCTTCTGGG
GGGGCCATGAGCATCTCAGGACTGATGATGTCACCCCTAGCTCGGGTCTTCCATCGGG
CATTTCCCAAGTGGCACCGCTTATTCACTAGACTTTCTACACTAGTAACCCACTGAAGATGG
CCAAGAAGGTTGGCCACCTGGCTGGATGCAACACACAGCACACAGTCTGGTAAACTGG
CTGAGGGCAGTATCAGGGCACCAAGGGATGATGCTGCTGTTCAACAGATGAGATTCTGG
GAACCTCCAGAGAGACCCGAAGAGATTATCTGGTCCATGAGCCCTGTGGGGATGGTGTGG
TGATGCCAGATGAGCTTGGTCTCTGACCCAGGGGAAGGTTCTATCTGCCCCACCTT
CTAGGTGTCACAAACATGGAAATTCAATTGGCTCTGCTTATAATACCAAGGGCAGGAGT
ACCACTTGTGGTGGAGGAGTACCTGGACAATGTCATGAGCATGACTGGAGATGTCAGAA
ACCGTAGTGGACATGAGTCAAGATGGCCTTCTGTGATGCCAACACTCGAGACTCTCAC
TACCCAGGAAAACCCAAATGATGGAAATCTGGGCTCTGGGACAGCTACAAACAGGATGAA
AAGTACCTCGAGCTGGATTATACCAACAGTGGCAT~~G~~AAGCTCAAGGAGAAGAAGATGG
TTTTGGATGAGTCTGTACAGTCTCAAAGACCTGAGAGACAGGCAATTCTAAGGGTGGC
TATGCAAGGAGGCCAAAGGGGGTCTACCCACCATCGAGGCCCTGGGGAGACTAGGCC
TGGACATACCTGGGACAGAGTTCTACCCACCATCGAGGCCCTGGGGAGACTAGGCC
GCCTCCAGGCCAAACTAGAGCTTGTGCTGGACCTGCACTGCCCTTTCAGGCC
TGACATCCCATGATGCCCTCTACTTCACTGTTGACATGCCAGTTAGGCCAGGCCCTGTCAAC
ACCAACTGTGTCAGCTCCAGGCCCTCAGGCCACCTTCTTCTTCTTCCAACTCT
CCCCACCTTCATGTCCTCTGTGACTCTTCTTATGGAGGCTGCCAGACTGCCACTG
CCCTGTCACTGCACCCAGCTGGCATTTACCATCCATCTGTCACACTTGTCTGTCTG
TCACATGGCTGGAGGCCCTAGGGCAGGTTGACATGGAGCAAACCTTTGGTAGTTGGG
TCTCTCTCCACCCACCATATCTCCCAAGGCCCTACCAAACTGATACACAGGGCTGG
TCTCTTCAATAAGAAGTGTGATAGAAAAAA

FIGURE 91

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179
<subunit 1 of 1, 545 aa, 1 stop
<MW: 58934, pI: 9.45, NX(S/T): 4
MSTGFSFGSGTGLGTTVAAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATTsapSSG
FGTGLFGSKPATGFTLGGTNTGALHTKRPQVVTKYGTLQGKQMHVGVKTPIQVFLGVPPSRPP
LGILRFAFPPEPPEPKGIRDATTYPPGW$LALS PGWSAVARSRLTATSASRVQASLLPQPLS
VWG YRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLNVYAPARAPGDPQLPVMWFPGGAF
IVGAASSYEGSDLAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIA
AFGGDPGNVTLFGQSAGAMSISGLMMSPLASGLPHRAISQSGTALFRLFITSNPLKVAKKVA
HLAGCNHNSTQILVNCLRALSGTKVMRVSNKMRFLQLNFQRDPPEEIIWSMSPVVVDGVVIPDD
PLVLLTQGVSSVPYLLGVNNLEFNWLPPYNITKEQVPLVVEYLDNVNEHDWKMLRNRMMD
IVQDATFVYATLQTAHYHRETPMMGICCPAGHATTRMKSTCSWILPQEWA
```

Important features:

Signal peptide:

amino acids 1-29

Carboxylesterases type-B serine active site.

amino acids 312-327

Carboxylesterases type-B signature 2.

amino acids 218-228

N-glycosylation sites.

amino acids 318-321, 380-383 and 465-468

FIGURE 92

GAGAACAGGGCTGTCTCAGGCAGGGCTTGCGCCCTCTATGCGGAGATGCTACTGCCACTGCT
GCTGTCTCGCTGCTGGCGGCTCCAGGCTATGGATGGAGATTCTGGATACCGATGCGAG
AGTCAGTGTGGTGCAGGGCTTGCACTCTGCTGCCTGCTCTTCTCTACCCCGA
CAAGACTGGACAGGGTCTACCCCGACTTATGCTACTGTTCAAAGCAGTGA
CTGAGACAACCAAGGGTGCCTGGCCACAAACCCACAGAGTGCAGAGGGTGGAAATGAGCACCCGGGGCC
GATTCAGCTCACTGGGATCCCGCCAAGGGGAACCTGCTCCTTGGTATCAGAGACCGCAG
ATGCAGGATGAGTCACAGTACTCTTCTGGTGGAGAGAGGAAGCTATGTGACATAATT
CATGAACGATGGTTCTTCTAAAGTAACAGTGCACGTTCAAGGGTGTGAGCGCACAGGAGGACC
ACAACACCGCCTACCTGGCATGGCTACAGGCTTCTCAGAGAAAGGGTGTGAGCGCACAGGAGGACC
GTCCGACTCCGTGCTTCTGGCTATGCCCGAGACCTTGTATCAGCATTCACGTGACAAACAC
GCCAGCCCTGGAGCCCCAGGCCAGGGAAATGCTCCATACCTGGAAGCCAAAAGGCCAGT
TCCCTGGGCTCTGTGCTGACGCCAGGCCACTGCTCAGCTGGGTCTCTGCAG
AAACAGAGTCCTCTCCCTGCTCCATCTGGGGCTAGACCCCTGGGCTGGAGCTGGCCGG
GGTGAAGGCTGGGATTCAAGGGCCTACACCTGCCAGGCCAGGAACAGGCTGGCTCCAGC
AGCGAGGCCCTGGACCTCTGTGCACTATCCTCCAGAACCTGAGAGTGTAGTGGTCTCCAA
GCAAAACAGGGCAGTCCTGGAAACCTTGGGAACGGCACGCTCTCCAGTACTGGAGGGCCA
AAAGCTGTGGCTGGTCTGTGTCACACACAGCAGCCCCCAGCAGGCTGAGCTGGACCCAGA
GGGGACAGGTTCTGAGCCCCCTCCAGCCCTCAGACCCCGGGTCTCTGGAGCTGCTGGGTT
AACTGGGAGCACGAAGGGAGACTTACCTGGCACTGCCAGCTGCCACCCACTGGGCTCCAGCAGT
CTCTCTCAGGGCTCTGGCACTATAAGAAGGGACTCATCTAACGGCATTTCTCAACGGAG
CGTTCTGGGAATCGGCATCACGGCTCTCTTCTCTGGCTGGCCCTGATCATCATGAAG
ATTCTACCGAAGAGACGGACTCAGACAGAAACCCCGAGGGCCAGGGTCTCCGGCAACAGCAC
GATCTGGGATTACATCAATGTGGTCCCAGCGGCTGGCCCTGGCTCAGAAGCGGAATCAGA
AAGCCACACAAAACAGTCCTGGGACCCCTCTCAGGGCTGGCTCAGGGCTGGGCTCAGAATCAAAG
AAAGAACCAAGAAAAGCAGTATCAGTTGCCAGTTTCCAGAACCCAAATCATCCACTCAAGC
CCCAAGAACCTCCAGGAGAGGCCAACGGAGCTCCATTATGCCAGCTCAACTTCCAGGCGTCA
GACCCAGGCTGAGGCCGGTAGGCCAAAGGCCAACGGCGATTATCAGAAGTCAGGTT
CAATGAGGGTCTTCTAGGTTTAGGACTGGACTTCGGCTAGGGAGGAAGGTAGAGTAAGAG
GTTGAAGATAACAGAGTGCAGGAAAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
CTCTCTCTCTCTCTCTTAAAAAACATCTGGGCAAGGGCAGATGCCAGGGTCTGAGCTCAGCCTGTAATC
CCAGCACTTGGGAGGGTTGAGGTGGGAGCTGCCAGTGGCTGAGGTGGGAGCTGGAGACGCCAG
GCCAACCTTGTGAAACCCCTCTACTAAACAAAATACAAAATTTAGCTGGCATGGTGCAGG
CGGGCTGTAACTCTACCTTGGGAAGCTGAGGCCAGGAAATCAGTGAACCTGGGAGACGG
AGGTTGAGCTGAGGCCAGGAACTCACCCATTGCAAGGCCAGCTGGGCAACAGGAGACTCCA
TCTCAAAAAAAAAACCTCCAAATGGGTTGGGTGTCGTAATCCAGCAGTCTGGAGGCTA
AGGTTGGGTGAGTTGCTTGAGGCCAGGAGTTGCAAGGCCAGCTGGGCAACATGGTGAACACCC
ATCTCTACAAAAAAATACAAAACATGCTGGCTTGGTGTGTCGCTGAGTCCAGCTGT
CAGACATTAAACCAAGGAACTCCATCTGGAAATAGGAGCTGAATAAAATGAGGCTGAGACC
TACTGGGCTGATTCTCAGACAGTGGAGGCATTCTAAGTCACAGGATGAGACAGGAGGCTCG
TACAAGATAACGGTCAATAAGAACACTTGGCTGATAAAAACAGATTGCAAGTAAAGAAGGCCACCAA
ATCCCAACAAAACCAAGTTGCCACGGAGGTGACCTCTGGCTGTCCTACTGCTACACTCC
GACAGCACCATGACAGTTACAAATGCCATGGCAACATCAGGAAGTTACCGATATGTCCCA
AAAGGGGGAGGAATGAAATAACCCACCCCTGTTAGCAAATAAGCAAGAAATAACCAAAAA
GTGGGCAACCAAGCAGCTAGGCCTGCTCTGTCTATGGAGTAGCATTCTTGTCCCT
TACTTCTTAATAAAACTGCTTACCTTAAACCA

FIGURE 93

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002
><subunit 1 of 1, 544 aa, 1 stop
><MW: 60268, pI: 9.53, NX(S/T): 3
MLPLLLLSSLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSFSYPRQDWTGSTPAYGYWFK
AVTETTKGAPVATNHQSREVEMSTRGRFQLTGDPAKGNCSLVIRDAQMQDESQYFFRVERGS
YVTYNFMNDGFFLKVTLSTPQDHNTDLTCHVDFSRKGVSQAQRTVRLRVAYAPRDLVIS
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLCAADSDQPPATLSWVLQNRLVLSSHPWGPRPL
GLELPGVKAGDSGRYTCRAENRLGSQQRALDLSVQYPPENLRVMVSQANRTVLENLGNGTSL
PVLEGQSLCLVCVTHSSPPARLSWTQRGVQLSPSQPSDPGVLELPRVQEHEGEFTCHARHP
LGSQHVSLSLVHYKKGLISTAFSNGAFLGIGITALLFLCLALIIMKILPKRRQTETPRPR
FSRHSTILDYINVVPTAGPLAQKRNQKATPNSPRTPPPAGSPESKKNQKKQYQLPSFPEP
KSSTQAPESQESQEELHYATLNFPGVRPRPEARMPKTQADYAEVKFQ
```

Important features:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 100-103, 297-300 and 306-309

Immunoglobulins and major histocompatibility complex proteins

signature.

amino acids 365-371

FIGURE 94

TGAAAGAGTAATAGTGGATACAAAGAGTCACCCAACTGAAGCTTATTCTGTGCTGCCTT
TATGTGGGAATTCTCTCTCATGGCCCTGTCTGGAGCAACAGAAAACCTCAACAAAGA
AAGTCACAGCAGCGACTGCAGATCTCATTTAGAGTCGAAGCGTGTGGGTGGAACCAATT
TTTGATCCAGAGGAAATGAATACGACTAGTCATCACATCGGCCAGCTAACATCTGATTTAGA
CAATGAAACAACTTCTTCAGTACAAGCTTGGGAGCTGGAGCTGGAAAGTACTTTTATCA
TTGAGAAAGAACAGGTGACATATATGCCATACAGAGCTTGTAGAGAGGGCGATCCCTC
TACATCTTAAAGAGCCAGGGATAATAGACATCTGCTACTGGAAGGGCTGTGGACACTGTG
GTTTGTCAAAAGTTCCGGATATCAATGACAATGAACAAAATTCTAGATGAACCTTATG
AGGCCATTGTACAGAGAGATGTCCTCAGAAGGAACATTAGTTATCCAGGTGACAGCAAGTGT
GCTGACGATCCCTCAAGTGGATAATATGCTGCTCTCTCACAGTACTTCAAGGCCAG
ATATTTCTGTGACCAAAACAGGACTCATAGAATATCTCTTAAATGGATAGAGAAC
TGCAAGATGAGTATTGGGTAACTCAAGCCAAGGACATGTTGGTCAGCCAGGAGCTTG
TCTGGAAACAAAGCTGTTATTTAAACATTCTCAGATGTTAACATGACAATAAGCCTTATTTAA
AGAAAAGTTTATACCCCTTGAATCTCTGAATCTGCAGGAACTCTCTGGAATTCTTACAGAACAA
TCATGGCATATGATAATGACATAGGAGAGAAATGCAGAAATGGATTACAGCATTGAAGAGGAT
GATTGCCAAACATTGACATTACTAATCATGAAACATCAAGAAGGAATGTTTATTTAA
AAAGAAAGTGGATTGAGCACCAGAGAACCTACTGGTATTAGAGCAAAGTTAAAACCATC
ATGTTCTGGAGCAGCTCATGAAGTACCAACTGAGGCTTCCACCACTTCTTACATAGATTCCAG
GTGGAAAGATGTTGATGAGGCTCTCTTTCTCTTCCATATTATGTTGAAAGTTTGA
AGAAAACCCACAGGGATCATTGTTAGGGCTGTGTCACAGACCCAGAACATAAGGAAAT
CTCCTATCAGTTCTTACTAGGAGCAAAGTGTCAATATCAATGATAATGGTAACTC
ACTACAAGTAACTCTGGATCTGAATACAGTCITGGTACAACCTAAGTATTACAGGCC
AGAAAAATAACAATATAGAACAGATCTCTCGATCCCCACTGTTGTCAGTCTTAAACATCA
ATGTCATGCTCTGGTACTCTCTCAATATCATGAGACTTATGTTGAAAGTCAGGCTCT
GGTCAGGTTAATCAGACTATCAGTCAGTGTAGAGATGAATCTCATAGAACAGCACCATT
TTACTTTAATCTATCTGAGAACACTAACATCAAGTTTACATCATAGATAATCAAG
ATAACACAGCTGTCATTTTGACTAATAGAACGTTTAACTCTCAAGAACCTGTTCTC
TACATCTCCATCTTAACTGGCCGAACTGGATCTCTGACTTCAACAAACCCCCTAC
CATCATCTGCTGTACTGGTGACAGTGGAGCACACAGACTGCCAGTACCCAGGAGCTTG
TGCTTTCCATGGGATTCAAGACAGAAGTTTACATGCTATTCTCATTTGATTATGATCATA
TTGGGGTTTATTTTTGACTTTGGGTTTAAACACAGGGAAACAGATTCTTCTG
GAAAATCTGAGAATTTCTCAGAGAAATATATCCTAACATGATGATGAGGGGGTGGAGAAAG
ATACAGAGGCCCTTGATATAGCAGAGCTGAGGAGTAGTACCATATGCGGGAACGCAAGACT
CGGAAACACAAAGCGTGAAGATCAGGGCCTATACAGGCACTTCTGCAAGTGGCCCGA
CAGTGGCCATTCTGGAAATTCTCAGGAGAACCTGAGGAGAACAGTAAACTGATCTGTG
CCCCCTCTTGTACTCTCTCAGACCTACAGGCTTCTGGAGAACAGGGCTTACAGTC
CTGAGCTCTTAAAGATCAGCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTAAAGGTT
GGGACCTCGCTTAAAGGATTAGCATGAGCTTGTGTCAGTGTGGAGCTCAAATAATTAGG
GCTTTTACCATCAAATTTTAAAGCTTAAAGTCTTACATGTTGATTCGACCCAACTGGTAGCTTAA
AGAGTTTGTGGCCCTGGCTCATGGCGGGAAAGCCTAGTCATGGAGTTTCTGATTCC
CTGGAGTAATACTCCATGGTTATTTAAGCTACCTACATGCTGTCTGGAAACAGAGATGTG
GGGAGAATGTAACAAACTCAGCTCACAGGCTACATCAACACAGGATGAGTAAATAATG
TAGGAAGATATAAAAGTAGATGAGAGGAGCACACAGATGTTAGCTGATCTTATG
CATTATTTACTTAGGAAAGAGTAAAAATACCAACAGGAAATTAAAGGAGCAAAATTG
CAAGTCAAATAGAAATGACAAATCGAGATAACATTACATTCTATCATATTGACATGAAA
ATTGAAAATGTATAGTCAGAGAAATTTCATGAAATTCTCATGAAAGTATTGTTCTTAT
TTAA

FIGURE 95

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906
><subunit 1 of 1, 772 aa, 1 stop
><MW: 87002, pI: 4.64, NX(S/T): 8
MNCYLLLRFMLGIPPLLWPCLGATENSQTKVKQPVRSRLVKRGWVWNQFFVPEEMNTTSHH
IGQLRSDLDDNGNNNSFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVSDINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARL
LYSLLQQPYFSVEPTTGVIRISSKMDRELQD BYWVIIQAKDMIGQPAGLSGTTSVLIKLSD
VNDNKPIFKESLYRLTVSEASPTGTSIGTIMAYDNDIGENAEMDYSIEEDDSQTTFDIITNHE
TQEGLIVILKKKDFEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLL
PYVFVFEETPQGSFVGVVSATDPNRKSPIRYSITRSKVFVNINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQITISAVDR
DESIEHHHFYFNLSVEDTNNSSFTIIDNQDNTAVILTRNTGFNLQEEPVFYISILIADNGIP
SLTSTNTLTIHVDCGDGSGSTQCQYQELVLSMGFKTEVIIAILICIMIIFGFIFLTLGLKQ
RRKQILFPEKSEDFRENIIFYQDDEGGGEEDTEAFDIAELRSSTIMRKTRKTTSAEIRSLY
RQSLQVGPDSAIFRKFILEEANTDPCAPPFDSSLQTYAFEGTGSILAGSLSSLESAVSDQD
ESYDYLNELGPRFKRLACMFGSAVQSNN
```

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518,
516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

FIGURE 96

ATTTCAAGGCCAGCCATATTTNTGTTGAACCAACAACAGGAGTCATAAGAATATTTNTA
AAATGGATAGAGAACTGCAAGATGAGTATTGGGTAAATCATTCAAGCCAAGGACATGATTGGT
CAGCCAGGAGCGTTGTTGAAACAACAAGTGATTAAATTAAACTTCAGATGTTAATGACAA
TAAGCCTATATTTAAAGAAAGTTATACCGCTTGACTGTTNTGAATCTGCACCCACTGGGA
NTTNTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTAC
AGCATTGAAGAGGATGATTGCAAACATTGACATTATT

FIGURE 97

GCAACCTCAGCTCTAGTATCCAGACTCCAGGCCGCCGGCGGGACCCCACCCGAC
CCAGAGCTTCAGCGGCCGCAGCGAGCAGGGCTCCCGCTTAACCTCCCGCGGGG
CCCAGCCACCTCGGGAGTCGGGTTGCCACCTGCAAACCTCCCGCTCTGCACCTGCCA
CCCCCTGAGGCCAGCGGCCGGGGGGGGAGCAGTCAGGCAACGCCGGGCTGCAGCTGGG
TTCAATTCTCCCTTCCTGGGATGGATCGGCCATCGTACGCCACTGCCCTGCCAGTGGA
TGTCTGCGTGTGCGAGAGCACCGGGAGATCCAGTGCAGGAAAGTCTTGACTCCTGCTGAAT
CTGAGCAGCACATTGCAAGCAACCGGTGATGGTGGCATCCTCTGGAGGTGAT
AGCAATCTTGTGGCCACGGTTGGCATGAAGTGTATGAAGTGTGGAAAGACGATGGTGC
AGAAGATGAGGATGGCTGTCAATTGGGGTGCATATTCTTCAGGCTGGCTATTITA
GTTGCCACAGCATGGTATGCCAATAGAAATCGTTCAAGAATTCTATGACCTATGACCCCAGT
CAATGCCAGGTACGAAATTGGTCAGGCCCTCTTCACTGGCTGGGCTGCTGCTCTCGCC
TTCTGGGAGGGGCCCTACTTGTGTTCTGCTGCCGGAAAACACCTTACCCAACACCA
AGGCCCTATCAAACCTGCACCTTCAGCGGGAAAGACTACGTGTAACAGAGGCAAAG
GAGAAATCATGGTAAACCAACCGGAAATGGACATTGAGATACTATCATTAACATTAGAC
CTTAAAGAATTGGTGTGAAATTGAACTTGTGATTGAGATACTAAACAAACAAACAAAAAA
ACCCATGTGTTAAAATACTAGTGTAAACATGCCCTAACTCTTATCTCTTCCCTCA
ATATAGGAGGGAAAGATTTTCATTGTATTACTGCTCCATTGAGTAATCATACTCAAAT
GGGGGAGGGGTGCTCTTAAATATATAGATGTGATATACATGTTTCTTAA
ATAGACAGTAAATACTATTCTATTGTGATGACTAGCATACTTAAATATCTCAA
AGGTAAATGTATTAAATCCATATTGATGAAGATGTTATTGGTATAATTCTTTCGTCC
TTATATACATGTAACAGTCAAATCATTTACTCTTCTTCAATTGAGTGTGCTGCCCTTGG
CCACAAGACCTAGCCTAATTACCAAGGATGAATTCTTCAATTCTCATGCGTCCCTT
CATACTTATTATTTTACCATTAATCTTATAGCACTTGCACTGTTATTAGCCCTT
TTGTTTGTGTTCAATTGCTCTATCTCTGAATCTAACACATTTCATAGCCTACATT
GTTCTAAAGCCAAGAAGAATTATTACCAAACTCAGAACATTGGAGGCAAATCTTGTGATG
ACCCAAAGTGATAAAATTCTGTGATGCCCTCCACAACTCCTGTACTCTGACCCATGACT
CTTGTGTTGCTTGGAAATATTGTCAATTGAGTAGCTGCATGCTGTTCCCCCAGGTGTTG
AACACAACCTTATTGATTGAAATTTTAAAGCTACTTATTCTATGTTTATATCCCCCTAAACT
ACCTTTTGTGCTCCCATTCCTTAATTGTGTTGTTCTCAAGGTGTAATTATCATGCTGTT
TATCTTCTAATAAGGTGGCTGTTCTGTTCTGTAACAAAGTGTAGACTTCTGGAGGTGATA
ATCTGTGTAACAAATTCTCTCTGTAGCTGTAGAAGCTTAATTCTTCTACCCCTTT
TTCTATCTGCGAAATTGAGATAATGATACTTAACTGCTGTTGAGAGGTAGTGTGAATTAA
TTAGTTTATTAATTACTCTTATTCTGAACTGAACATGACTATGCCCTATGTAATTGCT
CAGCTGGCTGAGACACTGAGAAGTCAGTGAACAAAACCTACACAGTACCTTCATGTGATT
CACTGCCCTTCTCTCTACAGTCTATTCCACTGAAACAAAACCTACACACATACCTTCT
GTGTTCTAGGCTCTCTCTCTCAACCTGCTATTCCACTGAAACAAAACCTACGCCACATAC
CTTCATGTGGCTCAGTGCCTTCTCTCTACCAGTCTATTCCATTCTTCAAGTGTGCT
GACATGTTTGTGCTCTGTTCAATTAAACAAACTGCTCTTCACTTTCCAGTCTGACAGAATG
CTATTCTACTTGAGCAAGGTGATGTAATTGGAAAGGGTGTGGCAGTGGTGTGGAGACCTG
GATTGAGTCTTGGTGTATCAATCACCGCTGTGTTGAGCAAGGGATTGGTGTGCTGTA
GCTTATTGCTTCATCTGTAAGCGGTGTTGTAATTCTGATCTTCCACCTCACAGTGTAT
TTGTGGGGGATCCAGTGTGAGATAACATGTAAGTGTGGGTTGTAATTAAAAAGTGTAT
ACTAAGGGAAAGAATTGAGGAAATTAACTGCATACGTTTGTGTTGCTTTCAATGTTGA
AAATAAAAAAAATGTTAAG

FIGURE 98

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185
><subunit 1 of 1, 211 aa, 1 stop
><MW: 22744, pI: 8.51, NX(S/T): 1
MANAGLQLLGFI LAFLGWIGAIVSTALPQWRIYSYAGDNIVTAQAMYEGLWMSCVSQSTGQI
QCKVFDSSLNLSSTLQATRALMVVGILLGVIAIFVATVGMKCMKCLEDDEVQKMRMAMIGGA
IFLLAGLAILVATAWYGNRIVQE FYDPMP TVNARYEFGQALFTGWAASLCLLGGALLCCSC
PRKTTSYPTPRPYPKPAPSSGKDYY
```

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-102, 118-142 and 161-187

N-glycosylation site.

amino acids 72-75

PMP-22 / EMP / MP20 family proteins

amino acids 70-111

ABC-2 type transport system integral membrane protein

amino acids 119-133

FIGURE 99

TTCTGGCCAAACCGGGGTCNAGCTGGCTTCACTCGCCTTCTGGGATGGATCGGC
GCCATCNCACACTGCCCTCCCCAGTGGAGGATTTACTCCCTATGCTGGGACAAACATCG
TGACCGCCAGCCCAGTACGAGGGGCTGTGGATGTCNCNGCGTGCAGAGCACCGGGCAG
ATCCAGTCAAAGTCTTGACTCCTGTGAATCTGAGCAGCACATTGCAAGCAACCGGTGC
CTTGATGGGGTTGGCATCCTCTGGGAGTGTAGCAATCTTGTGCCACCGTGGCATGAGTGTATGAAGTGTGGCATGAGCAGTGGGAGTGGCTGTCAATTGGGGGCGCAGATATTCTCTGGCAGGTCTGGTATGGCAACAGCATGGTATGGCAATAGAAN
CNTTCAACANTTCTATGACCCCTATGACCCAGTCATGCCAGGTACAATTGGTCA
GGCTCTCTTCACTGGCTGGCTGCTGCTTCTCTGCCCTCTGGGAGGTGCCCTACTTGTCA
GTTCTCTGTC

FIGURE 100

ACCCCTTGACCCAACGCGGCCCCCGACCGNTTCATGGCCAAACGCGGNCTCCAGCTGTTGG
GCTTCATTCTCCCCTTCCCTGGATGGACCGGCATCAGCACTGCCCTGCCCTGAGTG
GAGGATTACTCCTATNCCGCNAAACATCGTACCGCCCAGGCCNTGTACGAGGGCTGT
GGATGTCCTGCGTGTGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTGACTCCCTGCT
GAATCTGAGCAGCACATTGCAAGCAACCCTGCCTTGATGGTGGTTGGCATCCTCTGGAG
TGATAGCAATCTNNNTGGCCACCGTTGNNNTGAAGTGTATGAAGTGTGGAAAGACGATGA
GGTGCAGAAGATGAGGATGGCTGTCAATTGGGGCGCGATAATTCTTGCAGGTCTGGCTA
TTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCTATGACCGA

FIGURE 101

GGGCCCGACCATTATCCAACGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANC CGCG
CATCNCNTCAGACTCCCTGCCCATGGAGATTTNNCCTATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGGAGGGCTTGAACGTCNGCGTGTGCGAGANCACCGGCAGATCCAGTGCAA
AGTCTTGACTCCTTGTGAATCTGNGCAGCACATTGAGCAACCCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGGAGTGATAGCAATCTTGTGGCCACCGTTGGCATGAAGTGTATGAAGT
GCTTGGAAAGACGATGAGGTGAGAAGATGAGGATGGCTGTCAATTGGGGCGCGATATTCTT
CTTGCAGGTCTGGCTATTNNNNGTGCACAGCATGGTATGGCAATAGAACATGTTCAAGAAT
TCTATGACCCATTGACCCCAGTCATGCCAGGTACGAATTGGTCAGGCTCTCTTCACTGGC
TGGGCTGCTGTTCTCTGCCTCTGGAGGTGCCACTTTGCTGTTCTGC

FIGURE 102

ATTCTCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTCCCCANTGGAGGATTNAC
TCCCTATGCTGGCACAACATCGTGACCCCCAGGCCATTACGAGGGCTTGATGTCNT
GCNTGTCGAGCACCAGGAGATCCAGTGCAAAGTCTTGACTCCTTGCTGAATCTGAG
CAGCACATTGCAAGCAACCGTGCCTTGATGGGGTTGGCATCCTCTGGAGGTGATAGCAAC
CTTTGTGGCACCAGTGGCATGAAGTGTATGAAGTGCTTGGAAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCAATTGGGGGCGCGATATTCTTGTGCAGGTCTGGCTATTAGTNGC
CACAGCATGGTATGGCAATAGANTNNNTCNNGNNNTCTATGACCCATGACCCCAGTCAATG
CCAGGTACGAATTGGTCAGGCTCTTCACTGGCTGGCTGCTGCTTCTCTGCCTCTG
GGAGGTGCCACTTTGCTGTTCTGTCCC

FIGURE 103

AGAGCACCGCAGATCCCAGTNCAAAGTCTTGACCCCTGCTGAATCTGAGCAGCACATTNC
AAGCAACCCCTTGCCTTGAAGGTGGTTGNCATCCCCCTGGGAGTGAATAGCAATCTTGTG
GCCACC GTGGCATGAAGTNTATGAAGTGC TTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCA TTGGGGCGCGATATTTCTTCTGCAGGTCTGGCTATTAGTNNCCACAGCAT
GGTATGGCAATAGNATNNTCGNGGNTCTATGACCCCTATGACCC CAGTCAATGCCAGGTAC
GAATTTGGTCAGGCTCTCTCACTGGCTGGCTGCTGCTCTCTGCCTTCTGGGAGGTGC
CCTACTTTGCTGTT CCTGTCCCCGAA

FIGURE 104

AGCAATGCCCTGCCCCAGTGGAGGATTAATTCTATGNTGGGGACAACATTGTGACNGCCC
AGGCCATGTACGGGGGGCTGTGGATGTCCTGCGTGTGCAGAGCACCAGGAGATCCAGTGCA
AAAGTNNTTGACTCCTTGCTGAATTGAGCAGCACATTGCAAGCAACCGTGCCCTGATGGT
GGTTGGCATCTTCCCTGGGAGTGTAGCAATCTTGTGCCACCGTGGNAATGAAGTGTATGA
AGTGCCTTGGAAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCAATTGGGGCGCGATATTT
CTTNTTGCAGGTCTGGCTATTTAGTTGCCACAGCATGGTATGGCAATAGAATNGTTCAAGA
ATTTTATGACCCCTATGACCCAGTCATGCCAGGTACGAATTGGTCAGGCTTNTTCAC TG
GCTGGGCTGCTGCTTNTTCTGCCTNTGGAGGTGCCCTANTTGCTGTTCCCTGCGAAC

FIGURE 105

TCATAGGGGGCGCGATATTTTCTTGAGGTNTGGTTATTTAGTTGCCACAGCATGGTA
TGGCAATAGAATCGTTCAAGAATTNTATGACCCATGACCCAGTCAGTACGAAT
TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTNTGGGAGGTGCCCTA
CTTTGCTGTTCCCTG

FIGURE 106

TTCCTGGGATGGATCCGCCCATCNTCACATGCCCTGCCCNNTGGAGATTACNCCTATGC
TGGCGAACAAACATCNTGACCGCCCAGGCCATGTACGAGGGGCTGTGAATGTCCTGCGTGTC
CCAGAGCACCAGGGCAGATCCAGTGCAAAGTCTTGACTCCTTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCTGGGAGTGATAGCAATCTTG
CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGCGCGATATTCCTCTGCAGGTCTGGCTATTTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAAACTCTATGACCCCTATGACCCAGTCAATGCCAGGT
CGAATTGGTCAGGCTCTTCACTGGCTGGCTGCTGCTTCTCTGCCTCTGGGAGGT
CCCTACTTGCTGTTCTGTCCCCGAAAAACACCTTTACCCACG

FIGURE 107

CGGGGCTGCAGCTGTTGGCTTCATCTCGCTTCCCTGGGATGGAATCGGCGCCATCGTCAGCA
CTGCCCTGCCCATGGAGGATTTACTCNATGCTGGCGACAAACATCGTGACCNCAGGCCA
TGTACGAGGGCTGTGGATGTCNGCGTGTGAGCAGAGCACCGGGCAGATCCAGTGCAAAGTCT
TTGACTCCCTGCTGAATCTGAGCAGCACATTGCAAGCAACCNNTGCCTGATGGTGGTTGGCA
TCCTCCTGGGAGTGATAGCAATCTTGCGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTG
GAAGACGATGAGGTGAGAAGATGAGGATGGCTGTCATTGGGGCGCGATATTTCTTCTTG
AGGTCTGGCTATTTNTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTAT
GACCCCTATGACCCCAGTCAGTCAATGCCAGGTACGAATTGGTCAGGCTCTTCACTGGCTGGC
TGCTGCTTCTCTGCCTCTGGGAGGTGCCCTACTTGTGTTCTCGCAA

FIGURE 108

GGCTGCCGTAGCTCGCCGGCACCGGCCCTGCCCTCGCCCTGCCCTGCCCTGCGCTGCAC
CGCTGAGACGCCACCCCCCCTCAGCGGCCACCCGGTAGAGGCCCGCCCGTGCCCCG
ACCGGTTCCCCGCTTTGTAACCTTAAAGCGGGCGCAGCATTAACGCTTCCCCTGGG
GACCTCTAGGGTCTCCCGCCAAAGGTGCTGCCGGCTAAGAACATGCGCAAGGGAG
CAGGTCCCTGAGCTCGAGCGCAGCACCGACTCAAATTCCGAGGTCTTCACCGATGTTG
CACCAACCAACTAAAGCTGGCAACCCGACAGACCGAAATGTTGCTTTAAAGGTGAAGACTA
CAGCACCGACTGGTAGCTGAGGCCAACAGCGGAATCATCGATGCGAGGGCTCAATT
AATGTTATCTGTGATCTAACGCCCTTCCAGTATGATCCAATGAGAAAAGTAAACAAAGTT
TATGGTTCTAGTCTATGTTGCTCAACTGACACTTCAGATATGGAAGCAGTATGGAAGGAGG
CAAAACCCGAAAGCCTTATGGATCTAAACTTAAAGTGTGTTGAACTTCCAGCAGAGAAAT
GATAAAACACATGATGTTAGAAAATAAAATTATTCACAAACTGCTCAAAAGACGAAAC
ACCAATAGTGTCTAAGTCTCTGAGTTCTTGGATGACACCGAAGTTAAGAAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTAGGGCTACGGGAGGAAACAGCAGTTCAAG
GAAGAAGATGGATCTGCGATGAGGAACAGCTGCGAGCAACAGCCCATTTGAGCATTTG
CCCAACTGGAAAGGAAGGGCTTAGACCCGGCTCTGGCTTGTGGTTTGTCTTCTTGA
TCGTTGGTGTAAATTATTGGAAGATTGCCCTG**TAGAGGTAGCATGCACAGGATGGTAAATTG**
GATTGGTGTGATCCACCATATCATGGGATTTAAATTATCATAACCATGTTGAAAAGAAATT
AATGTTATGTGACATCTCACAGGTCTGGCTTAAATTACCCCTCTGCGACACACATCAC
AGATACACACACACAAATATAATGTAACGATCTTTAGAAAGTTAAAATGTTAGTAACGT
ATTGAGGGGAAAAAGAATGATCTTTATAATGACAAGGGAAACCATGAGTAATGCCCAAT
GGCATATTGTAATGTCATTTAACATGGCTAGGCTTGGTACATGATGCTGGATTACCTC
CTTAAATGACACCCCTTCCTGGCTTGGTCTGGCCCTGGGGAGCTGGAGCCAGCAT
CTGGGGAGCTGGCTCAGCTTCCACACAGTCTGGCCCTGGGGAGCTGGAGCCAGCTG
CTTCCGCTCTTCAGTTGCTGCAAGGCATCAGCTCTTGGGACTGATGAAACAGAGTCAGA
AGCCCAAGGAATGACTGTCAGCAGCATCACAGCTAATGTCATAATGAGGAGGGTGTG
TACTGATTGACCCAGCGCTTGGAAATAATGGCAGTGTCTTCACTTAAAGGGACCAA
GCTAAATTGATTGGTCTAGTAGTGAAGTCACACTGTTATTCAAGAGATGTTATGCTA
TTAACCTTTAAATGTTATCTCATGTTTCTTATTGTCACAAGAGTACAGTTAATG
TCTGGCTGCTGTAACCTGTTGGTAGCTGTTGCTGGGGCTGTTGGCTCTCT
GTCCTCTGGAGAGTCTGTCATGGAGGTGGGGTTATTGGGATGCTGGAGAACAGCTGCCA
GGAAGTGTTTTCTGGTCAGTAAATAACACTGTCATAGGGAGGAAATTCTCAGTAGT
ACAGTCACACTGAGTTACCTTTAAATGAGAGTAGCTGTCAGTCTCTAGTGTCTTATA
CCACCTCTCAACCATCTACACTTCACGGCCACGGGCTCAACTGTCAGCTGCCCT
TTGGGGACTAGCTGGAGTCAGGACAAATGGATCGGCTGCAAGGGTTAGAAGGGAGG
ACCAGCACTGTTGGTGGGGAGCAAGGGAGAGGAAACTCTCAGCGAATCTCTAGTAC
TAGTGTGAGAGTGTGACTGTAATTAAATTATGCTTAAAGGGACCAAACCCAGTCTGGT
CTATGATCATCTGGAAAAGAAAATTATAATAACGGCCAAAATTAGAAA

FIGURE 109

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977
<subunit 1 of 1, 243 aa, 1 stop
<MW: 27228, pI: 7.43, NX(S/T): 2
MAKVEQVLSLEPQHELKFRGPFTDVVTNLKLGPNPTDRNVCFKVTTAPRRYCVRPNSGIID
AGASINVSMQLQPFDYDPNEKSCHKFMVQSMFAPTDTSDEAVWKEAKPEDLMDSKLRCVFE
LPAENDKPHDVEINKIISTTASKTETPIVSKSLSSSLDDTEVKVMEECRKRLQGEVQLREE
NKQFKEEDGLRMRKTVQSNSPISALAPTGKEEGLSTRLLALVVLFIVGVIIGKIAL
```

Important features:

Transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

FIGURE 110

GTCAGTCTCTAGATTGTCCTTATCCCACCTTCAACCAACTCACATTTCNAGGCCAG
GTCCANGTCTGAGCCTGACTTCCCCCTGGGGACCTAGCCTGGAGTCAGGACAATGNTCGGG
CTGCAGAGGNTTAGAAGCGAGGGCACAGCAGTTGGTGGGAGCAAGGGNNAGAGAAA
CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTGACTGTGAATTAAATTATGCCATA
AAAGACNAACCCAGTTCTGTTGACTATGTAGCATCTGAAAAGAAAAATTATAATAAGCC
CCAAAATTAAGAATTCTTTGTCATTTGTCACATTTGCTCTATGGGGGAATTATTATTTT
ATCATTTTTATTATTTGCCATTGGAAGGTTAACCTTAAATGAGC

FIGURE 111

TATTGTAAAGGCCATTTAACCATGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNTTCCTCGCCTGTTGGTGCTGGCCTTGGGAGCTGGAGCCCCAGCATGCTG
GGGAGTGCAGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCGGCCCAGGCTGCTT
CCGTGTCTTCAGTTCTGTCAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAACCC
CAAAGGAATTGCCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCCTGTTGA
CTGATTGACCCAGCGCTTGGAAATAATGGCAGTGCTTGTTCACCTAAAGGGACCAAGCT
AAATTGTATTGGTTCATGTAGTGAAGTCAAACTGTTATTCAAGAGATGTTAATGCATATTTA
ACTTATTTAATGTATTCATCTCATGTTTCTTATTGTCACAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACCTGTTGGGTGAACGGTATTGCTGCTGGAGGGCTG

FIGURE 112

CCCTGGTGGTTTGTCTTTAATTCGTTGGTGTAAATTNTGGGAAGATTGCTTAGAGGTA
GNATGCACCNNGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTAT
CATAACCATGTGTAAAAAGAAATTAAATGTATGATGACATNTCACAGGTATTGCCTTAAATT
ACCCATCCCTGNANACACATACACAGATACACANANACAAATNTAATGTAACGATNTTTAG
AAAGTTAAAAATGTATAGTAAC

FIGURE 113

GGTGGCCCATCCCGGCCAGGCTGTTCCGGNTTCAGTTCTGTCAGGCCATCAGCTCC
TTGGGACTGTGAACAGAGTCAGAACGCCAAAGGAATTGCACTGTGGCACCATNAGACGTAC
TTGTNATAAGTGAGAGGGCTGTGTTGACTGATTGACCCAGCGCTTGGAAATAATGGCAGT
GCTTTGTTCANTTAAGGGACCAAGCTAAATTGTATTGGTTCATGTAGTGAACTGCAA
TTATTCAAGAGATGTTAACATGCATATTAAANTTATTAAATGTATTNATNTCATGTTTCTTA
TTGTCACAAGAGTACAGTTAACATGCTGGCTGCTGCTGAANTNTGTGGGTGAACTGGTATTGC
TGCTGGAGGGCTGTGGGCTCCTCTGTCITGGAGAGTCGGTCACTGTGGAGGTGG

FIGURE 114

TGCTTCCGTCTTCAGTTGTCCAAGCCATCAGCTCCTGGGACTTGATGAACAGAGTC
AGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTG
TGTTGACTGATTGACCCAGCGCTTGGAAATAATGCCAGTGCTTGTTCACTTAAAGGGAC
CAAGCTAAATTGTATTGTTCATGTAGTGAAAGTCAAACGTGTTATTCAAGAGATGTTAACATGC
ATATTTAACCTATTTAATGTATTTCATCTCATGTTTCTTATTGTCACAAGAGTACAGTTAA
TGCTGCGTGC

FIGURE 115

AAACCTTAAAAGTTGAGGGAAAAGAATGATCCTTATTAAATGACAAGGAAACNTGN
GT
AATGCCACAATGGCATATTGTAATGTCATTTAACATTGGTAGGCCTGGTACATGATGC
TGGATTACCTCTTAAATGACACCCTCCCTGCCTGTTGGTGTGGCCCTGGGAGCTN
GAGCCCAGCAGCTGGGGAGTGCGGTCTGCTCCACACAGTAGTCCCCANGTGGCCANTCCC
GGCCCAAGGCTGCTTCCGTCTCAGTCTGTCCAAGCCATCAGCTCCTGGGANTGATGA
ACAGAGTCAGAAGCCCAAAGGAATTGCANTGTGGCAGGCATCAGANGTANTNGTCATAAGTGA
GAGGCGTGTGTTGANTGATTGACCCAGCGCTTGGAAATAATGGCACTGCTTGTTCANTT
AAAGGGNCCAAGNTAAATTGTATTGGITCATGTAGTGAAGTCAAANTGTTATTGAGATG
TTTAATGCATATTAANTATTAAATGTATTCATNTCATGTTCTTATTGTCACAAGGGT
ACAGTTAATGCTGCGTGTGAANTCTGTTGGGTGAANTGGTATTGCTG

FIGURE 116

GGCCCTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGC GGTCAGCTCCACACAGTAGTCCC
CACGTGGCCCACTCCGGCCAGGCTGCTTCCGTCTTCAGTTCTGTCCAAGCCATCAGC
TCCCTGGGACTGATGAACAGAGTCAGAAAGCCAAAGGAATTGCACTGTGGCAGCAGACG
TACTCGTCATAAGTGAGAGGGGTGTTGACTGATTGACCCAGCGCTTGGAAATAATGGC
AGTGCTTGGTCACTTAAAGGGACCAAGCTAAATTGTATTGGTTCATGTAGTGAAGTCAAA
CTGTTATTCAAGAGATGTTAATGCATATTAACTTAAATTAAATGTATTTCATCTCATGTTTC
TTATTGTACAAGAGTACAGTTAATGCTGCGTGTGCTGAACTCTGTTGGGTGAACTGGTAT
TGCTGCTGGAGGGCTGTGGCTCCTCTGTCTCTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 117

GGAGCTCCGGGTGCTGGGCCGCTGGCGGGCGGCCTCGGCTCAGGCTGGCTGAGA
GGCTCCAGCTGCACTGGCTCCCGGCCGCTCTCGGAGCTCATCTGAGCTGAGTGCC
CTCGGGACAAACAGCTGGCAGGGTCTCATTTGTGCCCCAGGCTGGAGTCAGTCCA
TGATCATGGTTACTCGAGCTTGAACCTCTGGGTTCAAGGATCTGCTGAGTAGCTGGGA
CTACAGGACAAAATTAGAAGATCAAATGGAAATATGCTGCTTGGTTGATATTTCAC
CTGGGTGACCCCTATTGATGATCTGAAATGGAATGGATTTTATGTCGGACTATGAAA
GGTACCCCGATTGTCAGTGAAGGACTTCCATCTCACAGCCCCGATTGAGGAGATG
CTAACAGATGATGTTAAATACAGTGTGCACTGGAATGCCAGAGAAACTCCAACTCCCAGC
CTTTCTGAATTGGAGGATTATCTTCTATGAGACTGTTGAGAATGGCACCGAACCTT
AACAGGGTGAAGGTCAGATTTGGTTCTGGAGCCACTCAAAATATCACACCAAAAGGGAG
TATCTGTAGGAGAAAGAGCACAGGTGATGCCACCGACAGCAGTTGCACTTGGACA
AGGTTCTAACCAATTCCCTTCAGCACAGCTGTGAAGCTTCACGGGCTGAGTGGCAT
TCTCATTTCCCCTCAGCATGTTCTAACTGTCGCCACTGTGTTCATGATGAAAGGACTATG
TCAAAAGGGATAAAAGCTAACAGGTTGGGTTGTTGAAGATGAGGAAATAAGTGGAGGAAG
AAACCTGAGGTTCTAACAGGAGGAGCAGGAGAAGAGCTAGTGGTGTGACCAAGAGGGTAC
CAGAGAGCATCTGCAAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAAAATCTGGCCGGGTG
AGAGGATTCTGCCAGGGAGGGCTTCCAGTGGACCCGGGTAAGAAATCCCAATTCCG
AAGGGCTGGGACAGGGAGGCGATGGGAGCCTACCTGGATCTGGACTATGACTATGCTTCTGGA
GCTGAAGCTGCTCACAAAAGAAATACATGGAACTTGGAACTGCCCCAACGATCAAGAAA
TGCCCTGGTGAATGATCCACTTCAGGATTGATAACGATAGGGTGTACATTGGTCTAT
CGGTTTGCAGTGTGTCAGCACATCCAAATGATCTCTTAACTACTGGCATGTGACTG
GGGCTTCAACGGGTTGGGGGCTATCTGCTGTCAGAACATGCCAACGACAAAAGATTGGAA
GCAAAATCATGCGGTCACTCAGGGCACCACTGGGGATGTCACGGGGTTCAAGAGGAC
TACAACGTTGCTGTTCCATCCTCCCTAAACAGGCCAGATTGGCTCTGGATTACCGG
GAACAGTCCAAATTGCTTACGGCTAACAGGACACTGAAACAGGGCGGTATCATCTAAA
TCACAGAAAACACCAGCTGCTTACGGTACTAGTGAATCATCTTCAAGGTTATGGCTGACTT
GAACCTGTCAATGCAATTCAACATTTCAAAATCAGGAGATTTCGTCAATTAAAAAA
TGTATAGGTCGACAGATTGAAACATTAGGGCACTTCAATGCCAGTATATACTCTTCTTAA
CATGGTGATGAGTTTCTATGTTGAGAAAATTGTTGCTCTTCTAAAATAGACACACTT
AAACCTTCAACAGGTATTATAAAACATGTCAGTCTCTTAAATGGACTTATCTCAGGHTCC
TACTCTAACAGGAAATCTAATAGGATGCTGGTTGTATTAAATGTAAGGAAATTGCA
GGTAGATGGTAAAGGCAATTAGTATCAGAATAGAGACAGAAAGGTTCAACACAGTTGACTA
CTCTGAGATGGATCATTCATTGCTGCTGCCATGTCATGGCTTCAATTGTTGTTATCTGG
GGGACATTAGTTAGTTTGAAGAATTACAAATCAGAAGAAAAGCAAGCATTATAAA
AAAACAAATTAACATTGTTTACTGCTTAAAGAATAACAACTTACATGTTATTATAAAA
TGGGAAAGAAATGTTGTTCTATGAAATAACCTAGTTAGAAATTAGGAAAGCTGAGACATT
TAAGATCTCAAGTTTCTTAACTAATCTAAATATGGACATTCTCATGATGCTAGGG
AAGACACTTCACAAATTATGAATGATCATGTTGAAGGCCATTATTATTTATGCTATACAT
TCTATGATGAGGTTGCTCATTTTCTAGGACAAAGAAGATTCTGTAATCTTCTCAAGAAAAGACT
CTTTTCTCTGGACAAATTCAGCTTGTGAGGACTATAGGGTGAATTCTGTGATTAG
TAATTITAGATATGCTTCTTCTTAAATGAAATAAAATTATGAATGAA

FIGURE 118

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253
<subunit 1 of 1, 413 aa, 1 stop
<MW: 47070, pI: 9.92, NX(S/T): 3
MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFAEADAKMMVNVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLRVKQDLVLEPTQNITTKGVSVRRKRQV
YGTDSRFSIILDKRFLTNFPFSTAVKLSTGCGILISPQHVLTAAHCVHDGKDYGKGSKKLRV
GLLKMRNKS GGKKRRGSKRSRREASGGDQREGTREHLQERA KGGRRKKSGRGQR IAEGRPS
FQWTRVKNTHIPKGWAR GGMG DATLDYDY ALLELKRAHKKKYMELGISPTIKKMPGGMIHFS
GFDNDRADQLVYRFC SVSDESN DLLYQYCDAESGSTGSGVYRLKDPDKKNWKRKIIAVYSG
HQWVDVHG VQKDYNVA VRITPLKYAQICLWIHGNDANCAYG
```

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTAGCACCAGTACTGGAT
GTGACAGCAGGCCAGAGGAGCACTTAGCAGCTTATTCACTGTCGATTCTGATTCCGGCAAGG
ATCCAAGCATGAAATGCTGCCGTCGGCAACTCCTGGCACACTGCTCTCTTGCTTTG
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACGGGACGGCTATGGGATGCGCTG
GGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCTCTACTCTCTGAGGCCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAAATATCCGATACAGAACATGCACTGAGTAATGTG
TGCCCACCAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATATAATGATGTCAGCA
CCATGGCCAGTTTATGAATGGCTTCTGTGCTAATGACCCCTGACAACCCATGTTCACTCA
AGTCCAAGGAAAGGAACAACCCCTGGTTGTAAGTACGACCTAAGGCTTAGATGGTACG
CGTTGCTATACAGAATCTTGGATATGTGCACTCAGTGGTTATGCCAAATTGTTGGCTGCGA
TCACCCAGCTGGGAAGCACCGTCAGGAAGATAACTGTGGGTCTGCAACGGAGATGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAAATCCAGCTCCGCAACCAAATCGGATGATACT
GTGTTGCACTTCCATGGAAAGTAGACATATTGCGCTGTCTTAAAGGTCTGATCACTT
ATATCTGGAAACCAAAACCCCTCAGGGACTAAAGGTGAAAACAGTCAGCTCCACAGGA
CTTCTCTGGACAATTCTAGTGTGGACTTCCAGAAATTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTTCATTGCAAGATTGTAACTCGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTATCAACCCATCATCCACCGATGGAGGGAGACGGATTCTT
CTTGGCTCAGAACCTGTGGAGGAGGTTATCAGCTGACATGGCTGAGTGCTACGATCTGAGG
AGCAACCCGTGGTTGCTGACCAACTGTCACTATTACCCAGAGAACATCAAACCCAAACC
CAAGCTTCAGGAGTGCACACTTGGATCCTGTCCAGCCAGTGACGGATAACAGCAGATCATGC
CTTATGACCTCTACCATCCCCCTCTCGTGGAGGCCACCCATGGACCGCGTGCTCCCTCC
TCGGTGTGGGGGGGATCCAGAGCCGGCAGTTCTGTGTGGAGGAGGACATCCAGGGCA
TGTCACTTCAGTGGAAAGAGTGGAAATGCATGTCACCCCTAAGATGCCATCGCGAGCCCT
GCAACATTGGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGACAGTGCATGT
GGCCAGGGCCTCAGATACCGTGTGGCTCTGCATGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCCAAAACAAAGCCCCACATAAAAGAGGAATGCATGTAACCACTCCCTGCTATA
AACCCAAAGAGAAAATTCAGTCAGGGCAAGTTGCAACAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTGTCAGAGGAGGCCCTGTAAGTGTAAAAGCACAGACTGTTCTATA
TTTGAAGACTGTTTGTTAAAGAAAGCAGTGTCTACTGGTTGAGCTTCAAGGGTTCTGA
ACTAAGTGTAAATCATCTCACCAAGCTTTGGCTCTAAATTAAAGATTGATTAGTTCAA
AAAAAAA

FIGURE 120

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847
<subunit 1 of 1, 525 aa, 1 stop
<MW: 58416, pI: 6.62, NX(S/T): 1
MECRRATPGTLLLFLAFLSSRTARSEEDRDGLWDANGPWECSRTCGGGASYSLRRCLS
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVHHGQFYEWLPVSNDPDNPCSLKCQ
AKGTTLVVELAPKVLGDGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFL
VDNSSVDFQKFPDKEILRMAGPLTADFIKVIRNSGSADSTVQFIFYQPIIHRWRETDFFPCS
ATCGGGYQLTSAECYDLRSNRVVADQYCHYPENIKPKPKLQECNLDPASPQGYKQIMPYD
LYHPLPRWEATPWTACSSCGGIQSRAVSCVEEDIQGHVTSEEWKMVTPKMPIAQPCNI
FDCCPKWLAQEWSPTCTCGQLRYRVVLCIDHGMHTGGCSPKTAPHKEECIVPTPCYKPK
EKLPVEAKLPWFQQAQELEEGAAVSEEPS
```

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

FIGURE 121

CGGACGCGTGGCGCGCTGCGGAACCTCCGTGGAGGGCGGCTGGCCCTCGGCCCTGAC
AGATGGCAGTGGCCACTGCGCGCAGTACTGGCCCTCTGGCGGGCGCTGTGGCTGGCG
GCCCGCCGGTTCTGGGGCCCAGGGTCCAGCGGCTGCGCAGAGCGGGGACCCGCCAT
GCACGGGAAGACTGTGCTGATCACCGGGCGAACAGCGGCTGGCCGCGCACGCCCG
AGCTACTGCCCTGGAGCGCGGGTGTATGGGCTGCCGGACCGCGCGCGGCCAGGGAG
GCGCGGGGTCAAGCTCCGCCGAGCTCCGCCAGGCCGGAGTGCGGCCAGAGCTGGCGT
CAGCGGGGTGGCGAGCTCATAGTCGGAGCTGGACCTCGCCTCGCTGCGCTCGTGC
CCTTCTGCCAGGAAATGCTCCAGGAAGGCCCTAGGCTGGATGCTTGATCAATAACGCAGGG
ATCTTCCAGTGCCTTACATGAAGACTGAAGATGGGTTGAGATGCAAGTTGGAGTGAACCA
TCTGGGGCATTCTACTCACCAATCTCTCTGGACTCCTCAAAGTTCAAGCTCCAGCA
GGATTGTGGTAGTTCTTCAAACCTTATAAATACGGAGACATCAATTGATGACTTGAAC
AGTGAACAAAGCTATAATAAAAGCTTTGTTATAGCCGGAGCAAACGGCTAACATTCTTT
TACCAGGAAACTAGCCCGCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTTGAC
GTATTGTACGGACAATCTGGGGAGGCACATACACATTCCACTGTTGGTCAAACCAC
AATTTGGTGTACGGCTTTTCAAACCTCCAGTAGAAGGTGCCAGACTTCCATT
GGCCTCTCACGTGAGGTAGAAGGAGTGTAGGAAGATACTTTGGGATTGTAAGAGGAAG
AACTGTTGCCAAAGCTATGGATGAATCTGTTGCAAGAAAACCTGGGATATCAGTGAAGTG
ATGGTTGGCCTGCTAAAGAGGAACAAGGAGTAAAGAGCTGTTATAAAACTGCATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAATTGTTACTGAAAGAAAAAGAATT
ATATTGGAAATAGCCTGCTAACAGGGTACATGTGGGTATTTGGAGTTACTGAAAAA
TGGGATAAGAGAACATTCAAGCAAGATGTTAAATATATATAGTAAGTATAATGAATA
GTACAATGAAAAATACAATTATATTGAAATTATAACTGGGCAAGCATGGATGACATATTA
ATATTGTAGAACATTAAAGTGAACGGTACAGGTTTCAAGTATCTTGGTT
TCATGGCCAAAGTGTAACTAGTTACTACAATGTTGGTGTGGAAATTATCTGC
CTGGTGTGACACACAAGTCTTACTTGGAAATAATTACTGGTAC

FIGURE 122

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747
<subunit 1 of 1, 336 aa, 1 stop
<MW: 36865, pI: 9.15, NX(S/T): 2
MAVATAAAVLAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMCGRDRARAEAAGQLRRELQRQAAECGPEPGVSGVGELIVRELDLASLRSVRA
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNHLGHFLTNLLLGLKSSAPSR
IVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSRSKLANILFTRELARRLEGTNVTVNLHPG
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK
```

Important features:

Signal peptide:

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

FIGURE 123

GGGGATTGTAAGAGGAAGNACTGTGCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAATAGGAACAAAGGAGTAAAAGAGCTGTT
TATAAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTACT
TGAAGAAAAGAATTTGATATTGGAATAGCCTGNTAAGAGGNACATGTGGTATTTGGAG
TTACTGAAAAATTATTTTGGGATAAGAGAATTCAGCAAAGATGTTAAATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATAACAATTATATTGTAATTATAACTGGGCA
AGCATGGATGACATATTAATTTGTAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTT
TTCAAGTATCTTGAGTTCATGGCCAAAGTGTAACTAGTTACTACAATGTTGGTGT
TGTGTGAAATTATCTGCCTGGCTT

FIGURE 124

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCGGCTCCGGAGCCCAGCC
CTTCCCTAACCAACCAACCTAGCCCAGTCCCAGCCGCAGCGCCTGCCCTGTCAAGGGC
CCCAGCGTTACCAATGCATCCTGCCGCTGCCCTATCCTTACCCGACCTCAGATGTCCTCT
GCTCTGGTAACTGGGTTTTACTCCTGTAACAACAGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTAAACATGCTGATGTTGCTTAGTAAATTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTGGAGGAAGCTCCGATGTCATTAAGGAAGAA
TCCAAATGAAAATCAAGTAGTGTGCCCCAGAGTTGATTGATGTCAGCACTCTGACATAGGCC
AGAGATACAGGATAAGCAAATACCCAACCCCTCAAATTGTTGTAATGGGATGATGAGA
AGAGAATACAGGGGTCAGCGATCAGTGAAGCATTGGCAGATTACATCAGGCAACAAAAAG
TGACCCCATTCAAGAAATTGGGACTTAGCAGAAAATCACCCTCTGATCGCAGCAAAGAA
ATATCATTGGATATTGAGCAAAGGACTCGGACAACATAGAGTTTGAAAGAGTAGGG
AAATTTTGATGACTGCTTCTTCATGGGATGTTCAAAACCGGAAAG
ATATAGTGGCACAACATAATCACAACCCAGGGCATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTGATGACTTACAATTGGATTCAAGATAATGTTCTCTT
GTCCGAGAAATAACATTGAAAATGGAGAGGAAATTGACAGAAGAAGGACTGCCCTTCTCAT
ACTCTTCACATGAAAGAAGATAACAGAAAGTTAGAAATATTCCAGAAATGAAGTAGCTCGC
AAATTAAATAAGGAAAGGTTACAATAAACTTTTACATGCCGATTGTGACAAATTAGACAT
CCTCTCTGCACATACAGAAAACCTCCAGCAGATTGTCTGTAATCGTATTGACAGCTTAG
GCATATGTTGACTGCTGGAGACTTCAAAGATGTTAAATTCTGGAAACTCAAGCAATTG
TATTGACTTACATTCTGGAAACTGCACAGAGAATTCCATCATGGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGGCCAAGATGTAGCAAGCAGTCCACCTGAGAGGCTCTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTTATTGAGGGATCGAGATGAGCTTAAACTTG
AAAAACAGTTGTAAGCCTTCAACAGCAGCATCAACCTACGTGGGAAATGTAACCTA
TATTTCATAATTCTATGTTGTTTATTGAAATAACAGAAAGAAATTAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 125

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689
<subunit 1 of 1, 406 aa, 1 stop
<MW: 46927, pI: 5.21, NX(S/T): 0
MHPAVFLSLPDLRCSLLLLTVWVFTPVTTEITSLATEINIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMKREYR
GQRSVKALADYIRQQKSDPIQEIRDLAETTLDRSKRNIIGYFEQKDSNDNYRVFERVANILH
DDCAFLSAFGDVSKPERYSQDNIIFYKPPGHSAPDMVYLGMATNFDVITYNWIQDKCVPVLVREI
TFENGEEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFRHMVFGDFKDVLIPGKLKQFVFDLHSGKLHREFHHGPDPDTAPG
EQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL
```

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

FIGURE 126

ATTAAGGAAGAATTCCAATGAAAATCAAGTAGTNTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATAACAGGATAAGCAAATACCCAACCCTCAAATTGTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGTCAGCGATCAGTGAAAGCATTGGCAGATTA

09473514.101601

FIGURE 127

AGAGGCCTCTGGAAAGTTGTCGGGTGTTGCCGCGNGAGCCGGTCAAGAGGGACNAGG
TGCGCTGCCGGAGAACCTCCTCCGCTGCCGCTCCGGAGCCCAGCCCTTCTTAACCC
AACCCAAACCTAGCCCNGTCCCAGCCGCCAGCGCTGTCCTGTCNCGGANCCAGCGTNACC
ATGCATCTGCCGTCTCCTATCCTAACCGACCTCAGATGCTCCCTCTGCTCTGGTAAC
TTGGGTTTTACTCCTGTAACAACGTAAATAACNNGTCTTGATACNNAGAATATAGATGAAA
TTTTAACNATGCTGATGTGGCTTAGTCATTTATGCTGACTGGTGTGTTCAAGTCAG
ATGTGGCATCCAATTTGAGGANGCTCCGATGTCATTAAGGAAGAATTCCAAATGAAAA
TCAAGTAGTGTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA
TAAGCAAATAACCAACCCCTCAAATTGTTGTAATGGGATGATGATGAAGAGAGAATACAGG
GGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGC

FIGURE 128

GCCCCACGCGTCCGATGGCGTTCACGTTCGCGGCCCTCTGCTACATGCTGGCGCTGCTGCTCA
CTGCGCGCCTCATCTTCTTCGCCATTGGCACATTAGCATTGATGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCAGTGTAACTCCCTGAATCCCCTGTACTCCAGAGTACCTCAT
CCACGCTTCTTCTGTGTCATGTTCTTGCGAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCCCTTGGCATATCATATTGGAGGGTATAGAGTAGACCAAGTGTGAGTGGCCAGGA
CTCTATGACCCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG
GTGCAAATTAGCTTTTATCTTCTAGCATTTTTACTACCTATATGGCATGATCTATGTTT
TGGTGAGCTCTTAGAACACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAGCCAC
CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
TACTTTAAAAAAATGACTCCTTATTTTAAATGTTCCACATTTTGCTGTGAAAGACTG
TTTCATATGTTAACTCAGATAAGATTAAATGGTATTACGTATAAAATTAAATATAAAAT
GATTACCTCTGGTGTGACAGGTTGAACCTGCACTCTTAAGGAACAGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCTAGTACATTGGAAAGCTTTGTTATAGGAACATTGTA
GGGCTCATTGGTTTCATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
TTCTGATGAAGTGAAAATGATATCTGACTAGTGGAAACTTCATGGGTTCCCTCATCTGTC
ATGTCGATGATTATATGGATACATTACAAAAAATAAAAGCGGGAATTTCCTCGCTT
GAATATTATCCCTGTATATTGATGAATGAGAGATTTCCCATATTCATCAGAGTAATAAA
TATACTTGCTTAATTCTAACGATAAGTAAACATGATAATAAAATATGCTGAATTACTT
GTGAAGAATGCATTAAAGCTATTAAATGTTTATTGTAAGACATTACTTATTAAAG
AAATTGGTTATTATGCTTACTGTTCTAACTGGTGGTAAAGGTATTCTTAAGAATTGCAAGG
TACTACAGATTTCAAAACTGAATGAGAGAAAATTGTATAACCCTGCTGTTCCCTTAGT
GCAATACAATAAAACTCTGAAATTAAAGACTC

FIGURE 129

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330
<subunit 1 of 1, 144 aa, 1 stop
<MW: 16699, pI: 5.60, NX(S/T): 0
MAFTFAACFCYMLALLLTAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHAFF
CVMFLCAAEWLTLGLNMPPLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS
```

Important features:

Signal peptide:

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

09978504 101601

FIGURE 130

ATTATAGCATTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCACTGTAAATACCCCTG
AATCCCCTTGTACTCCCAGAGTACCTCATCCACGCCCTCTTCTGTGTATGTTTCTTGTC
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCTCTGGCATATCATATTTGGAGGTATA
TGAGTAGACCAGTGATGAGTGGCCAGGACTCTATGACCCTACAACCATCATGAATGCAGAT
ATTCTAGCATATTGTCAGAAGGAAGGGATGGTCAAATTAGCTTTTATCTTAGCATT
TTACTACCTATATGGCATGATCTATGTTGGTGAGCTCTTAGAACACACAGAAGAATT
GGTCCAGTTAAGTGCATGCAAAAAGCCACCAATGAAGGGATTCTATCCAGCAAGATCCTGT
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTTAAAAAATG

FIGURE 131

CGGACGCGTGGGGAAACCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAACAAGATGCGCGCGCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCGCCG
TGCTGCTGCTGACCATGGCCTTGGCGGAGGTCGGGACCGCTTCGGCTGAAGCATTTGAC
TCGGTCTTGGGTGATACGGCGTCTGCCAACGGGCCTGTCAGTTGACCTACCCCTTGACAC
CTACCTAAGGAAGAGGAGGTGACGATGTCAAGAGAGGTTGCAGGCTGTTCAATTGTC
AGTTGTGGATGATGGAATTGACTAAATCGAACTAAATTGAAATGTAATCTGCATGTACA
GAAGCATATTCCAATCTGATGAGCAATATGCTGCCATCTGGTGCAGAACATCAGCTGCC
ATTGCTGAAGACAAGAACAACTTATGCTCCTGATGCCAAAATGCACCTACTCTTTC
CTCTAACTCTGGAGGTGACATGATGGACTCCGCACAGAGCTTCATAACC
TCTTCATGGACTTTTATCTCAAGCCGATGACGGAAAAATAGTTATTCAGCTAAGCC
AGAAATCCAGACGCAACCATTTGGAGCAGGAGCCTACAAATTGAGAGAACATCTCTAA
GCCAAAATGCTCTATGCAAATGAGAAATTCAAGCGCACAGGAATTCTGAAAGATGGA
GAAAGTGAATGGCTTTTAAGATGCTCTCTTAACTCTGGGTGGATTAACTACAACCTCT
TGTCTCTCGGTGATGGTATTGCTTGGATTGTTGTCGAACTGTTGCTACAGCTGTGGAGC
AGTATGTTCCCTGAGAACAGCTGAGTATCTATGGTGAATTGGAGTTATGAATGAAACAAAG
CTAACAGATATCCAGCTTCTCTTGTTGTTAGATCTAAACTGAAGATCATGAAGA
AGCAGGGCCTCACCTACAAAAGTAATCTGCTCATCTGAAATTTAAGCATTTC
AAAAGACAAGTGTAAATAGACATCTAAATTCCACTCTCATAGAGCTTTAAATGGTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAATAAGTTACTCAAATCTGTG

FIGURE 132

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847
<subunit 1 of 1, 323 aa, 1 stop
<MW: 36223, pI: 5.06, NX(S/T): 1
MAAPKGSLWVRTQLGLPPLLLTMALAGGSGTASAEAFDSVLGDTASCHRACQLTYPLHTYP
KEEEELYACQRGCRLFSICQFVDDGIDILMRTKLECESACTEAYSQSDEQYACHLGCONQLPFA
ELRQEQLMSLMPKMHLLFPLTLVRSFWSDMMDSAQSFITSSWTFYLQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSYLQMRNSQAHRNFLEDGESDGFLRCLSLNSGWILTTLVL
SVMVLLWICCATTAVATEQYVPSEKLSIYGDLFMNEQKLNRYPASSLVVRSKTEDHEEAG
PLPTKVNLAHSEI
```

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

FIGURE 133

TTGGGTGATA CGGC GTCTGCCACCGGCCTGTCAGTTGACCTACCCCTTGACACCTACCC
TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTCAATTGTCAAGTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTGGTTGCCAGAACATCAGCTGCCATTGCG
TGAAC TGAGACAAGAACAACTTATGTCCCTGATGCCAAAATGCACCTACTCTTCCTCTAA
CTCTGGTGAGGTCAATTCTGGAGTGACATGATGGACTCCGC

FIGURE 134

CACACTGGCCGGATCTT TAGAGTCCTTGACCTTGACCAAGGGTCNGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGGAACAAAGATGGCGGCCGAAGGGAGCCTTTGGGTGAGGAC
CCAACTGGGCCTCCCGCGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCGGGACCG
CTTCGGCTGAAGCATTGACTCGGTCTGGGTGATAACGGCTTGGCCACCGGGCTGTCAG
TTGACCTACCCCTTGACACCTAACCTAAGGAAGAGGGATTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTCAATTGTCAGTTGTGGATGATGGAATTGACTTAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAACATATTCCAATCTGATGAGCAATATGCTTGCATCTT
GGTGCCAGAATCAGCTGCCATTGCTGAAC TGAGACAAGAACAACTTATGTCCTGATGCC
AAAAATGCACCTACTCTT CCTCTAAC TCTGGT GAGGT CATTCTGGAGTGACATGATGGACT
CCGC

FIGURE 135

GCAGGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCTGGGAGGCAGGCCGGAGGT
GGGGCGCCGCTGGGCGGCCGACGGCCTCATCTGAGGGCGACGGCCCGAACGGAGC
GTGCGGACTGGCCTCCAAAGCTGGGCGACAAGCTGCCGGAGCTGCAATGGGCGCGCTG
GGGATTCTGTTGGCCTCTGGCGCCGTGGCTGCTCAGCTGGCACGGAGAGGAGC
AGCCCCCGGAGACAGCGGACAGAGGTGCTCTGCCAGGTTAGTGGTTACTGGATGATTG
ACCTGTGATGTTAAACCATTGATAGATTAAACTACAGGCTTCCAAAGACTACAAAA
ACTCTTGAAGTGAACACTTTAGGTATTACAAGGTAACCTGAAGAGGCCGTGTCCTTCT
GGAATGACATCGCCAGTGTGGAAGAAGGGACTGTGCTGTCAAACCATGTCATCTGATGAA
GTTCTGATGGAATTAAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA
AGAATGTGAACAAGCTGAACGACTGGAGCAGTGGATGAATCTCTGAGTGAGGAAACAGA
AGGCTGTTCTCAGTGGACCAAGCATGATGATTCTCAGATAACTCTGTAAGCTGATGAC
ATTCACTCCCCTGAAGCTGAATATGTAGATTGCTCTTAATCCTGAGCGCTACACTGGTTA
CAAGGGACCAAGATGCTGGAAAATATGGAATGTCATCTACGAAGAAAACTGTTTAAGCCAC
AGACAATTAAAGACCTTAAATCCTTGGCTCTGGTCAAGGGACAAGTGAGAGAACACT
TTTCACTAGTTGGCTAGAAGGTCTGTGTAGAAAAAAAGAGCATTCTACAGACTTATATCTGG
CCTACATGCAAGCATTAAATGCTTGTGAGTGCAAGATATCTTCAAGAGACCTGGTTAG
AAAAGAAATGGGACACAAACATTACAGAATTCAACAGCAGTGGATGGAATTGACTGAA
GGAGAAGGTTCAAGAAGGTTAAGAACCTGTATTCTACTTAATAGAACTAAGGGTTT
ATCCAAAGTGTACCATTCTCGAGCGCCAGATTCAACTCTTACTGGAAATAAAATTC
AGGATGAGGAAACAAATGTTACTCTGGAAACTCTCATGAAATCAAGTCATTCTGTTG
CATTTGATGAGAATTCTATTGCTGGGGATAAAAAGAAGCAACAAACAAAGGAGGA
CTTTCGACTGCATTAGAAATTCTCAAGAATTATGGATTGTGTTGGTTATTAAATGTC
GTCTGTGGGAAAGCTTCAAGACTCAGGGTTGGCACTGCTCTGAGATCTTATTTCTGAG
AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAAATTCCATCTAACAGACAAGA
AATAGTATCATTATTCAACGCATTGGAAAGATTCTACAAGTGTGAAAGAATTAGAAAAC
TCAGGAACCTGTTACAGAATTCTCATAAAGAAACAAAGCTGATATGTGCCGTGTTCTGGAC
AATGGAGGCGAAAGAGTGGATTCTCAAGGCATAATGCAATGACAGTCTTAAGCCAA
ACATTTTATAAAAGTGTGTTGAAAGGAGAATTATATTGTTAAGTAAACACATT
AAAAATTGTGTTAAGTCTATGTATAACTACTGTGAGTAAGTAATCTTAATAATGTG
GTACAAATTAAAGTTAATATTGAATAAAAGGAGGATTCAAAATTAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

FIGURE 136

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974
<subunit 1 of 1, 468 aa, 1 stop
<MW: 54393, pI: 5.63, NX(S/T): 2
MGRGWGFLFGLLLGVWLSSGHGEEQPPETAAQRCFCQVSGYLDDCTCDVETIDRFNNYRLF
PRLQKLLES DYFRRYYKVNLRKPCPFWNNDISQCGRRDCAVKPCQSDEVPDGIKSASYKYSEEA
NNLIEECEQAERLGAVDESLESETQKAVLQWTKHDDSSDNFCEADDIQSPEA EYVDLLNPE
RYTGYKGPDWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLQLQETWLEKKWGHNITEFQQRFDGILTEGEGPRLKNLYFLYLI
ELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLEILHEIKSFPLHF DENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFKCR LWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFH
LTRQEIVS LFAFGRISTSVKELENFRNLLQNIH
```

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAAACTGTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTGCTCTGGCAAGGGACAAGTGAAGAGNACACTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAAGGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGCAAGATATCTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAATTNAACAGCGATTTGATGGAATTTGACTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACCTGTATTTCTACTTAATAGAACTAAGGGCTTATCCAAAGTGTACCAATT
CTTNGAGCGCCCAGATTTCAACTNTTTACTGGAAATAAAATTCAAGGATGAGGNAAACAAAA
TGTTACTTTGGAAATACTTCATGAAATCAAGTCATTCCCTTGATTTGATGAGAATTCA
TTTTTTGCTG

FIGURE 138

CGGACGCGTGGCGGACGCGTGGCGGACGCGTGGGTGGGAGGGGCAGGATGGGAGGGAA
AGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTCTCATACTGGACAGAAC
CGATCAGGCATGAACTCCCCTTCGTCACTCACCTGTTCTGCCCTGGTGTTCCTGACAGG
TCTCTGCTCCCCCTTAACCTGGATGAACATCACCCACGCCATTCCAGGGCCACCAGAAG
CTGAATTGGATAACAGTGTCTAACATGTTGGGGTGGACAGCGATGGATGCTGGTGGC
GCCCTGGATGGGCCTTCAGGGACGGGAGGGGGACGTTATCGCTGCCCTGTAGGGGG
GGCCCACAATGCCCATGTGCCAAGGGCACTTAGGTGACTACCAACTGGGAAATTCATCTC
ATCCTGCTGTGAATATGCACCTGGGATGTCTGTAGAGACAGATGGTGATGGGGATT
ATGGTAGCTAAGGAGAGGGTGGTGGCAGTGTCTGAAGGTCCATAAAAGAAAAAGAGAA
GTGTGGTAAGGGAAATGGTCTGTGTGGAGGGTCAAGGAGTTAAAACCTAGAAAGCAA
AGGTAGGTAAGTCAGGGAGTAGTCTTCATGCCCTTCACACTGGGAGCATGTTCTGAGGGT
GCCCTCCCAGCCTGGAGTAACTATTTCCCCCATCCCAAGGCCCTGTGCCCTCTGGTCT
CGTCTTGTGGCAGCTGTCTTCAGTCTGGGATATGTGCCCTGTGGATGCTTCATTCCA
GCCTCAGGGAAGCCTGGCACCCACTGCCCAACGTGAGGCCAGAGGAAGGCTGAGTACTGGT
CCCAGAAGGAGATACTGGGAGGGAAAGATGGGCAAGCGGTATGATGCCTGGCAAAGGG
CCTGCATGGTATCCTCATTGCTACCTAATGTGCTGCAAAGCTCCATGTTCTCTAACAGA
TTCAGACTCTGGCCAGGTGTGGCCACCTGTAATTCTAGCATTGGGGAGGGCCAA
GTGGGCAGTCATTGGAGGTCAGGAGGTCAAGACCCAGCCCTGGCCACATGGTGAAACTCCAT
CTCTACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
ATCTACTCGGGGGAGGGCTAAAGACAGGGAGACCTCTCACTCAACCCAGGGGTGGAGGTGGGTGG
AGCCAAAGATGTGCCTGTCACTACGTAGCGGTGGGTGACAAGGTAAGCGAGGACTCCATCTCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
CTCATCGCCTGTAAATCCCAACATTTGGAGGGCCAACGCAGGAGGTGGTTGTAGGGCTGGGAG
GTTTGAGACCACGCCCTGGCAACATGAAAAGACCCCACTCTAAAAATGTTTAAAAAAATGTTTAAAAAA

FIGURE 139

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039
><subunit 1 of 1, 124 aa, 1 stop
><MW: 13352, pI: 5.99, NX(S/T): 1
MELPFVTHLFLPLVFLTGLCSPFNLDEHPRLFPGPEAEFGYSVLQHVGGGQRWMLVGAPW
DGPSGDRRGDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETDGDGGFMVS
```

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

0000733444 101601

FIGURE 140

CACAGTTCCCCACCATCACTCNTPCCATTCCCTTCCAACCTTATTTAGCTTGCCATTGGGA
GGGGGCAGGATGGGAGGGAAAGTGAAGAAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAACCGATCAGGCATGGAACCTCCCTTCGTCACTCACCTGTTCTTG
CCCCCTGGTGTCTGACAGGTCTCTGCTCCCCCTTAACCTGGATGAACATCACCCACGCCT
ATTCCCAGGGCCACCAGAACGCTGAATTGGATACAGTGTCTTACAACATGTTGGGGTGGAC
AGCGATGGATGCTGGTGGCGCCCCCTGGGATGGGCTTCAGGCACCGGAGGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGCCACAATGCCCATGTGCCAAGGGCCACTTAGGTGACTA
CCAACCTGGGAAATTCTCATCCTGCTGTGAATATGCACCTGGGATGTCTCTGTTAGAGA
CAGATGGTGTGATGG

FIGURE 141

AAAGTTACATTTCTCTGGAACTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGGGTGCTCGAGCCCCCTTCTGAGCTTCCTGGCCGCTCTAGAACAA
ATTCAAGGCTCGCTGCAGCTCAGACCTCAGCTCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAAATGCTTATTTGAAAGAAACATGTTCTAGGTCAAACGTGAGTCACCA
AATGCAGACTTTCAAAATGGTTCTAGAAGAAATCTGGAAACAAGTCTTTCTATGTGGTTTCT
ACGCAATTGATTCCATGTTGCTCACAGATAAGTGGCCATTCTGCTGCCCCCTCAGAACCTC
TCTGTAECTCAACCAACATGAAGCATCTCTGATGTTGGAGGCCAGTGTATCGGCCTGGAGA
AACAGTGTACTATTCGTGAAATACCAGGGGAATCAGAGGACCTGTACAGGACACATCT
GGATCCCAGCAGCTGGTCTCACTCTGAAGGGTCTGAGTGTGATGTCACTGTGACATC
ACGGCACTGTGCCAACATTCCTGTCAGGGCCACATTGGCTCACAGACCTGAGCCTG
GAGCATCTGAAGCATCCCTTAATAGAAAATCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGCTTCAACCTGGTTATGGTGGAGGACCTGGGCCAGTGGTTGAGTT
CTTGCGCTACTGGAGGGAGGCCCTGTGCCAGGAAACATGTCAAAATGGTGAAGGAGTGG
GGGTATTCCAGTGACCTAGAACCATGGAGCCAGGGCTGCAACTGTGTGAAGGCCAGA
CATTCTGTGAAGGCCATTGGAGGTACAGCGCCTTCAACGAGAACAGAAATGTGTGGAGGTGCAA
GGAGAGGCCACTTCTGGTACTGGCCCTGTTGCTTGTGCTTCTCATGCTGATCTTGT
GGTCTGCCACTGTTCGTCTGAAAATGGCCGCTGTCCAGTACTCTGTTGCCCGTGG
TGGTCTCCAGACACCTGAAAATAACCAATTCCCCAGAAGTTAACGCTGAGAAGG
GAGGAGGTGGATGCCCTGCAAGGGCTGTGATCTCTGAGGAAACTCTCAGGGCTGGAT
CTCATAGGTTGCGGAAGGCCAGGTGAAGCCGAGAACCTGGTCTGATGACATGAAACC
ATGAGGGGACAAGTGTGTTCTGTTTCCGCCAGGACAAGGGATGAGAGAACAGTAGGAAGA
GCCCTGTTGCTCACAAAGTCTAGAAGCAACCATCAGGGCAGGGTGGTTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGAAAAGTGACTTCATCCCTCGGTCTCAAGTTTCTCATCTGTAATGGGGGAAATTACC
TACACACCTGCTAAACACACACACAGAGTCTCTCTATATACACAGTACACATAAA
TACACCCAGCAGCTTGCAGGCCACTAGAGGGAAACTCTGTGACACTCTACAGTCTGACTGATTGAG
TGTTTCTGGAGAGCAGGCCAGGACATAATGTATGAGAGAACAGTCAAGGACTCTACACACTGGGT
GGCTTGGAGAGGCCACTTCCAGAATAATCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCAAGGCCAATGCCCTGAGAGGGGAATGGCTTACGGAGCTACAGT
AGGTGACCTGGAGGAAGGTACAGGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACACTACTGAAAGTGTGACAGTGTGACACTGAGCACAGCAGGGTGAATGTATGT
GTGCAATGCCAGAGAACATGAGAAGTCAGTAAACATGTGCAATTGTTGTTGCTCCCTTTTC
TGTTGGTAAAGTACAGAAATCAGCAAATAAAAGGGCACCCCTGGCAAAAGCGGTAAAAAA
AAAAAAAAAA

FIGURE 142

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033
<subunit 1 of 1, 311 aa, 1 stop
<MW: 35076, pI: 5.04, NX(S/T): 2
MQTFTMVLEEIWTSFLMWFYALIPCLLTDEVAILPAPQNLSQLSTNMKHLLMWSPVIAPGE
TVVYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSGQTSAW
SILKHPFPNRNSTILTRPGMEITKDGFLVIELEDLGPQFEFLVAYWRREPGAEEHVKMVRSG
GIPVHLETMPEGAAVCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLAFLFAFVGFMLILV
VVPLFVWKMGRLLQYSCCPVVLPDTLKINTSPQKLISCRREEVDACATAVMSPEELLRAWIS
```

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

FIGURE 143

TCTCGCTGATGCACATCTGGGTTGGCAAAGGAGGGTGCCTCGAGCCGCCCTTCTAGCTT
CCTGGCCGGCTCTAGAACAAATTCAAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAAGATGGCTGAGATGACAGAATGCTTATTGGAAAGAAACAATGTTCTAGG
TCAAAACTGAGTCTACCAAATGCAAGACTTACAATGGTTCTAGAAGAAATCTGACAAGTCT
TTTCATGTGTTTCTACGCAITGATTCCATGTTGTCACAGATGAAGTGGCCATTCTGC
CTGCCCCCTCAGAACCTCTGTACTCTAACCAACATGAAGCATCTTGATGTGGAGGCCA
GTGATCGCGCTGGAGAACAGTGTACTTCTGTGAAATACCCAGGGGGACTACGAGAGCCT
GTACACGAGGCCACATCTGGATCCCCAGCAGCTGGTCTCACTCACTGAAGGCTCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTGTGTCAGGGCCACATTGGC
TCACAGACCTCAGGCTGGAGCATCTGAAGCATCCCTTAATAGAAACTCAACCATCCCTAC
CCGACCTGGGATGGAGATCACAAAGATGGCTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCCAGTTGAGTTCCTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGTATCTCTGAGAAAAGAGAGGCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTGAGTGGGCTGGGTGAT

FIGURE 144

CCACCGCGTCGCCACCGCTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGAGGTGGGAGAGCACGAAGCTGGAGGCCACACTGAGGGAGGGCGGG
GGAGGTGAAGAAGGAGAGAGGGAGAAGAGGCCAGGAGCTGAAAGGAGAGAGGGAGGAGG
GAGGAGATGCCATGGAGACCTGGAGTTAGGTGGCTTGGGAGGCTTAATGAAAAGAGAAC
GGAGGAGGGTGTGGGTTAGAACCAAGAGGTAGCCCTGTGGGAGCAGAAGGCTGAGGAGA
GTTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTCCGGAAAAGAGCAGAGGAAAGAGG
AAAGACACAGAGAGACGGGAGAGAGAAGAGACTGGGTTGAAGGGGGATCTCAGTCCTG
GCTGCTTGGCATGGGAACTGGACTCCCTGTGGGAGGAGAGGAAAGCTGGAAGTCCT
GGAGGGACAGGGTCCCAGAAGGAGGGACAGAGGAGCTGAGAGAGGGGGCAGGGCTTGGG
CAGGGTCCCTCGGAGGCCCTGGGATGGGCTGCAGCTCGTCTGAGCGCCCTCGAGC
GCTGGTACTCTGGCTGCACTGGGGCAGCACGTCACATCGGACCACGACCTGACCCGAGG
ACTGGTGGAGCTACAAGGATAATCTCAGGGAAACTTCTGCCCCCTTCTGCCCCCATTAAAGGCTCAGCACTGGAGGAGAGA
AGCTCCGGGAACCTTGACAAACCGGGCAGATGTCCTCTGCCCTGCACCCGACCT
GTGCTCAATGTGCTGGAGTCCCTCTTACAGCCACCGACTCAGTGAACCTGCGCTGCT
GTTGGAGCTCGCAGGGAGCCGGCTCGAACATCAGATCAACCACCAAGGGCTTCTGCTG
AGGTGAGCTCATTCACTCAACCAGGAACCTACGGAAATTCTAGCGCTGCCCTCCCGGGC
CCCAATGGCCTGGCATTCTCAGCCTCTTGTCAACGTTGCCAGTACCTCTAACCCATTCC
CAGTCGCCTCTTAACCGCAGACCCATCACTCGCATCTCTACAAGAATGATGCCACTTT
TTCAAGACCTGAGCCTGGAGCTCTGTTCCCTGAATCTCTGGCTTCATCACCTATCAGGGC
TCTCTCAGCACCCCGCCCTGCTCCGAGACTGTCACCTGGATCCTCATTGACCGGGCCCTCAA
TATCACCTCCCTCAGATGCACTCCCTGAGACTCCTGAGGCCAGAATCCTCCATCTCAGATCT
TCCAGAGCCTCAGCGGTAAAGCCGGCCCTGAGCCCTTGGCCACAGGGCACTGAGGGC
AACAGGGACCCCGGCACCCCGAGAGGCCAGGCTGCCAGGCCAACACTACCGCCTGCATGTGGA
TGGTGTCCCCCATGGTCGTGA GACTCCCTCGAGGATTGACCCGCCCCCTAAGCCTC
CCCACAAGGGCAGGGAGTTACCCCTAAACAAAGCTATTAAAGGGACAGAAACTTAA

FIGURE 145

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353
<subunit 1 of 1, 328 aa, 1 stop
<MW: 36238, pI: 9.90, NX(S/T): 3
MGAAARLSAPRALVLWAALGAAAHHGPAPDPEDWWSYKDNLQGNFVPGPFFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPL
LYSHRLSELRLLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAAASRGPNGLAILSL
FVNVASTSNPFLSRLLNRTDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSTPPCSE
TVTWILIDRALNITSLQMHSRLLSQNPPSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER
RCRGPNYRLHVGVPHGR
```

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

FIGURE 146

GGCGCTGGTTCTGGGTACTGGCTGACGGAGCAGGGAGCAAGGGTCGCCAGCCCTCCGCCCGAGGCC
GTCCTGTCCTCCCCCCCCTCGCTCTCTGCAGCTACTGCTCAGAAACGCTGGGGGCCACCCCTGCGAGACTAACGAA
GCAGCTCCCTCCCCAACCCTGCAGGTCTAATTTGGACGGCTTGGCTCTGCAGACTAACGAA
CGAGGGCGGAGGCTCGCTTAITCTGCAGCAGCACCCAGCTGCGCCCGAGCCTGGGTCTCAGGGCCCTC
GGAGGGGGCTCTCGCTCTGGTGAAGGCTCTGGGGCTGAGAGCCGGCTCGGGTTGGCT
CACCTCTCCAGGAACAACTTCAACTGGAGGAGCCAAGGAGTGGAGAGGAGCTGCTTGGAGATTTCCTGGGAA
ATCTGAGGTCAATTCAATTGAAGTGTACCGCGGGAGTGCTCAGTAACACAGTCGTGTCATGGCTAGA
GCAATTCCAGCATGGTGGCTCCAACTGCACCTTATTGGAGAACATTGGAGGAAATCTTGGAGGATGT
GAGTGGTGGATAGCCAAACAGGGAAAGGAGCATACAGCAGACTGAGACTATTGGACCTTCA
AAATAATTCAAGAAGTCAGGTGTATCAACAGCCTTAATATGGAGTATATGACATGGGTGTAGAGCTGGAAAGA
TCAGCAAGTCTGGGGAACTGGGGAGTGGCTTGGAGAACATGGACCTGGTCTGGAGGAT
TTGGGAGCAGACTGGGAAGATAAGGGGGAGCTTGTACATCAATCTGGTATGATGAAGTGAAGACTTT
AGCTACCCATATGAACATGAAGCAACCCATATTGCTCAGTGGCTGCTGGCCCTGTATGACACATTATACA
CAGGTGTTGGGAACTTAGTAAGCAAGTCTGGTGTGCAATTACTCCCAAAGGAAACTGGTGGGGCATGCCCTAACAA
CATGGGGCCCTGGTGTGCAACCTAGTTTGAGGGGCTGTAGAGAAAATCTGTGAAAGAAGGG
TCAGACAGTATTATCCCCCTCGAGAGGAAACATGAAGATAGCAGCAGTCAGCAGTCAACAGTCAACACC
CACTGGGGAAAGATCAGATGAGTGGAGCAGAAATGAGTCAGCAGCAAGGAAATTGCTCCAAATTTG
TGTAAGATTAGAATAGGATCAGCAAGGAAACACTGCAATGGTACAGTGTACATTGCTGGTGTGGAT
AGTAAAGCTAAAGTTATGGCACTGTACATTATGAAATGCACTCAGCTGTAGAGCTGCAATTCAATTATGGT
ATAATAAGACATAATGATGGTGGGGTAGATATCAGTCAAGGAGAAGAGCATTATTCTCATGCAAC
AAAGTGTATTCAACAACTTGGCAATTATCAGTCTGCTTAATTCTTCAGTCTCAAGTAAAGTCAAGCTGTG
ACTTGTGAAACAACTGTGGACAGCTGTGCTTCAATTAACTGCTGCTCACTGGCCAAAGGATTAAGTCT
CGTAACTGTATGCAAGGAAATCACCATTATGCTGTGTAATTGGAAACTCGAGTTATTCTGTCTGCTAGTATC
TGCAGAGCACCAGTACATGCTGGACTGGTCAATCAGGGTGTATGATGTAAATGCCGTGGACAAAGA
AAAGACCTACATTGCTCTTCTGCAAGGAAATTCTTCAGAAAGGAGGACCTAAAGACTATTCAAAATGCAATTCTGA
ATTGGTATAAAACTGTAAACATTACTGTCAGAGTACATCAACTATTCTGCCCCAAAGGGTGCCTAAATGCTA
AAATCTGTATAACAAAGTCATAAAATAAACAGGGAAATTCTGGTGTGTAATTGGAAACTTCAACATTGATG
TTTGTGAAATCTGTGTTAAATATGCTATAATTCTTAGCAGTTATTCTACAGTTAAATTCAATAGCTGATT
GTCTCAGTTCTATATAATTATGCTGTTTGTATATGCCACTAAATAAAATGAATCTTCAACATTGATGTAATG
GCCCTGAAAAATCATCTAGTGCATTTAAATAATCAGACTCTAAACTGTAAAGAAGAACCTTACACATT
AGTTCATGCTATGCCATTACCAACTCCAAATACTCAAAATAATTCTCAACTTAATGTAAGGTTTTTC
TGTTAAATTAGGCATATAGAAATTAAATTCTGTGATATTGCACTCTTATTATATAATTCTTAAATC
CAAAATGAACTCTGTTAAATGTTGATTCTTGGGAATGCCCTTAAAGGAAATGTAATAAAAGTCAAGGTGGTGGT
ATGAAAACATTCTAGTGATCATGTCAGTAAATGTAGGGTTAACATGGACAGCCAGAGCTTCTATGTA
AAATTGAGGTGACATATAATTCTTGTATCTGGCAAAACTCTCTGAGGGAGGTATAATTGACAAAGTT
GAACAAAAGATGAACTTAATGTTACATACCAATTGCTGCTTCAACTGTTTAAATGTTAAATGACCTGTTGATATAA
ATATTGCGATATCATGGTACCTTAAATGCTGATATAATTGTTCTATGAAAAATGTATTTGCTTGTGTTGATACTAAA
AATCTGTAAGGTTAGTTGGTAAATTCTGGTGTGAAATTACATATTAAATTCTGGTGTGAA
TAAACATTTAAATCATGTTCAAAAAAA

FIGURE 147

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417
<subunit 1 of 1, 500 aa, 1 stop
<MW: 56888, pI: 8.53, NX(S/T): 2
MKCTAREWLRTTVLFLMARAIPAMVVPNATLLEKLLBKYMDEDGEWWIAKQRGKRAITNDNM
QSILDHLNKLRSQVYPTASNMEYMTDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHW
GRYRPPTFHVQSWYDEVKDFSYPSYHECNPYCPFRCSGPVCYHTQVVWTSNRIGCAINLC
HNMMIWQIWPKAVALVCNYSPKGNNWWGHAPYKHGRPCSACPPSFGGGCRENLCYKEGSDRY
YPPREETNEIERQQSQVHIDHTVRTRSDSSRNEVISAAQMQSIVSCEVRLRDQCKGTTCNR
YECPAGCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDNDGGWVDITRQGRKHYFIKSNRNGI
QTIGKYQANSFTVKVTQAVTCETTVEQLCPFPHKPASHCPRVYCPRNQCMQANPHYARVIG
TRVYSDLSSICRAAVHAGVVRNHGGYDVMPVDKRKTYIASFQNGIFSESLNQPPGGKAFRV
FAVV
```

Important features:

Signal peptide:

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77.

N-glycosylation site

amino acids 38-31

FIGURE 148

GGGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGAAGGCCGCCCCGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCCGGGCCGCCTCGACGGGCAGGCCCTCCCATGTCCCTGCTCCACGCCG
CGCCCTCCGGTCAGCATGAGGCTCTGGCGGCCGCGTGCCTGCTGCTGCTGGCGCTGT
ACACCAGCGCTGGTGGACGGTCAAATGCAAGTGCTCCGGAAGGGACCCAAGATCCGCTAC
AGCGACGTGAAGAAGCTGGAATGAAGCCAAGTACCCGACTGCGAGGAGAAGATGGTTAT
CATCACCAACCAAGAGCGTGTCCAGGTACCGAGGTCACTGCTGCACCCCAAGCTGC
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCCTGGAACGAGAAGCGCAGGGTCTACGAA
GAAAGGGTGAAAAACCTCAGAAGGGAAAACTCCAAACCCAGTGGGAGACTTGTGCAAAGGA
CTTTCAGATTAaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaAGCCTTC
TTTCTCACAGGCATAAGACACAAATTATATTTGTTATGAAGCATTACCAACGGTCAG
TTTTTACATTATAGCTGCGTGCAGAACGGCTTCCAGATGGGAGACCCATCTCTCTGTGCT
CCAGACTTCATCACAGGCTGCTTTTATCAAAAGGGAAAACTCATGCCCTTCCTTTAA
AAAATGCTTTTTGTATTTGCCATCGCACTACATCTGAGCTTTAAGCGCCGGGA
GGAACAATGAGCTGGTGGACACATTCAATTGCAGAGTGGTGCCATTCTAGCTGGGAAGC
TTCGCTTAGAGGTCTGGCCCTCGCACAGCTGCCACGGCTCTCTGGCTTATGCCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCTGTAGCCGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTGAGGAACTCAAGTTGGTGCAGAAAATGTGCTTCACTCCCCCT
GGTTAATTTCACACCCCTAGGAACATTCCAAGATCCTGTGATGGCAGAACAAAATGATC
CTTAAAGAAGGTGTGGGGTCTTCCCAACCTGAGGATTCTGAAAGGTTCACAGGTTCAATA
TTTAATGCTTCAAGACATGTGAGGTTCCAACACTGTCAGCAAAACCTTAGGAGAAAA
TAAAAATATATGAATACATGCCAATACACAGCTACAGACACACATTCTGTTGACAGGGAA
AACCTTCAAAAGCATGTTCTTCCCTCACCACAACAGAACATGCAGTACTAAAGCAATATAT
TTGTGATTCCCCATGTAATTCTCAATGTTAAAACAGTGCAGTCCCTTCGAAAGGTAAGAT
GACCATGCGCCCTTCCTGTACATATACCCCTTAAGAACGCCCTCCACACACTGCC
CAGTATATGCCGATTGTACTGCTGTTATATGCTATGTACATGTCAGAAAACCATTAGCAT
TGCATGCAGGTTCATATTCTTCAAGATGGAAGTAATAAAAATATTTGAAAGTAAAA
AAAAAAAAAA

FIGURE 149

MSLLPRAAPPVSMRLLAALLLLALYARVDGSKCKCSRKGPKIRYSDVKKLEMKPYPH
CEEKMVIITTKSVSRYRGQEHLHPKLQSTKRFIKWYNNAWNEKRRVYEE

Signal sequence:

amino acids 1-34

04978544-101501

FIGURE 150

GGCCCCAGGGACTGCTATGGCTTCCCTTGTGTTCACCCCGCTGGCTCATGTTAACTCCAATGCTCCCTGTG
GTTAACTGCTCTGGCATCAAGTCACCCCTCATGGCAGCCAGCACAGTATCCAGTGTCAACACAATTATGG
CAAACCGGAGGCCCTAAAGAACACCGTACCCAGATCTGGGTCAGGGAGCAGTACTTAAAGGGTCCCTA
TGCCCTACCCCCCCTAGGGAGAGGGCGTTTACGGCCCGAGAACCCCGCTCCTGGACTGGCATCCGAATAC
TACTCAGTTGTGCTGTGTCGCCCCAGCACCTGGATGAGAGATCCTACTGTCATGACATGCTGCCCATCTGGTT
TACCGCCAAATTGGATACTTGTGACTGCTATGTCAGATCAAAATGAAGACTGCCTTTACTTAAACATCTACGT
GCCCAAGGAAGATGGAGCCAAACACAAGAACAGATGATATAACAGATTAATGACCGTGGTGAAGACGAAGA
TATTGATGATCAGAACAGTAAGAGGGCTGATGGCTTATACATGGGGATCTTACATGGAGGGACCCGGCAA
CATGATTGACGGCAGCAATTGGGCAAGCTGGAAAGCTGTGATCCTGGATCATCCACCATTAACCGTGTGGGAATACT
AGGGTTTAAAGTACCGGTGACCAGGCAACAAAGGAACATATGGCTCTGGATCATGATTCAGACTGCGGTG
GATTGAGGGAGAATGTGGGAGCCCAAGAGATGACTTGGCTCGGGGCTGGGGCTGGGGCTC
CTGTTGTCAGGCTGTGACCCACTGTCCAACTCATGAGGGTCTCTTCCAGAGGGCATTCAGAGCGCACCCTC
CTGTCAGGCTGGGAGCTAACAGGCTGGGAGCCAGTACCTGGGAGACAGAAGGGTGGCTGCAACAT
GCTGGACACCCAGGATGGTAGATGGCTGGGAGCTGGGAGACAGAAGGGTGGCTGCAACATCCCC
GCCACCTACACCATAGCCTTGGCGGGTGTGATGACGGGAGCTCATCCAGACGACCCCCAGATCTGTGATGGA
GCAAGGGAGGTTCTCAGGACATCATGGCTGGGAGCTAACAGGAGGAGGGCTGAAGGACTCTGGACG
CTGGATATAACQGGAGCGGTGTGACGGCCACQACTCTGGACTTCTCTGGTCAACTCTGTGAGGCAACACTTACGG
CTACCCCTGAAGGGAAAGGACACTTGGGAGAGATCATGTCACAGACTGGGCCGATAAGGAAAACCC
GGAGAGCGCCGGAAACACCTGGGTGTTCTTACTGACCAAGCTGGGAGCCCGTGGCCCGCAGACCT
GCACGCGCAGTACGGCTCCACCTACTCTATGCTTCTATCATCAGCCTAACAGCAATGAAGGCCAGTG
GCTGGACATGGGAGCTTGGTAGGTTCTGGCCCTATGCTTGGCAGTCCACCCGATGATGGTCTCCACCGAGCTTCTAG
TGTAACCTTCCAAAGAACAGCCTATGCTGGCGTGTGATCATGGCAGACTTGGACAACTTCCCAAAACTGG
TGCCAAAGTATAACACAGCTCTATGTCATGGCATATTGGCTTGAACACAGATGAGAGATACTACCGGG
AACGAAAGTGGCTTCTGGTAGGACTCTGGTCTCTTGGACATGGCTTCTCATGGACACATCTGGAG
AACACCAAAGGGTCTCCACAGGACATGACATCATTTCCCTATGGCACCCGGATCTCCCGCAAGATATGGCC
AACACCAAAGGGCAGCAATCATCTCTGGCCAACTCCAAACACTTAAAGGACCTCTAACAAACAGGGCTGA
GGCACCAACTGTCTCATGIGAACACAGGATATTCCACCGAAATAAGGTGTCACCATGGCGTCGGGGCTC
GCTCTCTTCTCCAAACATCTTAACTGGCTTGGGGCGCTGTACTACAAAAGGACAAAGGGCCTGAGACATCACAG
GGGCCCCAGTCCCAGAGAAACCCACAAATGATCTGGCTACATCCAGAACGAAGGAGATCATGTCCTGGAG
GAAGCAGCTGGAACACGATCACAGTGAGTCGCTGCAGGACACGACACACTTGAGGCTCACCTGGCCGAGA
CTACCCCTCAGGCTGGCCGGTGGCCAGATGACATCCACATTATGGCAACACCATCACCAGTATTCAA
CACACTGACGGGGATGACGGCTTGGCAGCTTGGACACCTTTAACCCCTCAGTGAGGAGCAAACAGTACAAATTACCCA
CGGACATCTCCCAACTAGAGTATAGCTTGGCTTCTTCCCTTCTATCCCTCTGGCTTACCGCCTGAGCAACAT
AGAAAGGGGAAAGGAAAGAGAGAGGGAGGAGAGGAGAGGAGGAGAAAGTCTCCAGACGGGAATGTTTTGTCCTAC
GACTTAAAGCAAAATGCAAAAAGGAGCTGATCCCATCCCGAGAACCCCTTATGTTGGTGTGTTTCCAGTATTAC
AAGATCAACTCTGGACCTGTGAAATGAGAGATACACATTCTGTTAAATAACTGCTTAAAGATCTTACCA
CTCCAAATGATTTAGTGTGAGGACATCACCAATTCAAGGGCCGGGTGTTCCACGCTCATGGAGGAGCT
GACACTTGTGAAACTCAGGCCAAGGACACTGATATTTTAAATACAGGAACTTAAACATTCTTCTGTG
CACACAATGGATGGCTCTCCATTAGTGAAGAAAGAGTCAATGAGATTGGCCAGCACATGGAGCTGTAATCCAG
AGAGAGGGAAACGTAAGAAATTATTATTAAAGAATGGACTGTGCGAGCAGAAATCTGTACGGTTCTGTGCAAAG
GTGTTTGGCCAGCTGAACTTAAAGTGTGAAACTTGGCCAGCACATGGAGCTGTAATCCAG

FIGURE 151

MLNSNVLLWLALAIKFTLIDSQAQYPVVNTNYGKIRGLRTPLPNEILGPVEQYLGVPYASP
PTGERRFQPPEPPSSWTGIRNTTQFAAVCPHQHDLERSSLHDMPLIWFATANLDTLMTYVQDQN
EDCLYLNIYVPTTEDGANTKKNADDITSNDRGEDEDEDIHQNNSKKPVMVYIHGGSYMEGTGNI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAGNYGLLDQIQALRWEENVGAFFGDPKR
VTIFGSGAGASCVSLLTLSHYSEGLFQKAIIQSGTALSSWAVNYQPAKYTRILADKVGNCNML
DTTDMVECLRKNYKELIQQTITPATYHIAFGVIDGDPVIPDDPQILMEQGEFLNYDIMLGV
NQGEGLKFVDGIVDNEDGVTNFDFSVSNFVDNLYGYPEGKDTLRETIKFMYTDWADKENP
ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMPSWADSAHGEDEVPYV
FGIPMIGPTELFSCNFSKNDVMLSAVVMTYWTNFAKTGDPNQPVPQDTKFIHTKPNRFEVA
WSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHNLNEIFQYVSTTTKVPPDMTS
FPYGTRRSPAKIWPTTKRPAITPANNPKHSKDPMKTGPDFTVLIETKRDYSTELSVTIAVG
ASLLFLNILAFAALYYKKDKRRHETHRRSPQRNTTNDIAHIQNEEIMSLQMKQLEHDHECE
SLQAHDTLRLTCPDYTLTLRRSPDDIPLMTPNTITMIPNTLTGMQPLHTFNTFSGGQNSTN
LPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

04978544-101601

FIGURE 152

GGGAAAGATGGCGGCACTCTGGGACCCCTTGGGTCGGCAGCTGGCGGCATGTTGT
CGGCTGGGATGGTCCAGGATGTTACTCTTCTTGTGGGCTGGGCAGGGCCA
CAGCAAGTCGGGGCGGTCAAACGTTAGTACTTGAACAGGGAGCACTCGTGTGAAAGCC
CTACAGGGTGTGGGCACAGGCAGTTCCTCACTGTGAATCTGATGGCAATGCCATGGTGA
TGACCCAGTATATCCGCTTACCCAGATATGCAAAAGAACAGGGTGCCTTGTGAAACGGG
GTGCCATGTTCTGAGAGACTGGGAGTGGCTGGCACTTCAAATCCATGGACAAGAAA
GAAGAATCTGCATGGGATGGCTGGCAATCTGGTACACAAAGGATGGATGCAGGCCAGGG
CTGTTGGAAACATGGGAAATTCTGGCTGGGAGTTTGAGACACCTACCCCAAT
GAGGAGAACGAGCAAGGAGGGTATTCCCTACATCTCAGGCATGGTGAACAAACGGCTCCCT
CAGCTATGATCATGAGCGGGATGGCGGCCTACAGACTGGGAGGCTGCACAGGCATGTCC
GCAATCTCATTAGACACCTCTGGTGAATTGCTACAGTCAGAGGCATTGACATAATG
ATGGATATTGATGGCAAGTGGAGGGACTGCATTGAAGTGGCCGGAGTCCGCTGCC
CCGGCTACTACTTCGGCACCTCCCTCATCTACTGGGATCTCTCAGATAATCATGATGTC
TTTCTTGAAGTTGTTGAACTGACAGTGGAGAGAACCCCAGAAGAGGAAAGCTCCATCGA
GATGTTCTTGGCCCTAGGGACAAATTAAGCTGCTGGAGATGACAGCTCCACTGCC
CCTGAGTGGCTGGGCTTCCCTCATGCTCTTCTCTCTGGTTCTGTATTGCA
TAGTCATTGGTACATACCTACAAACATGGCAGGAACAGGCCAAAGCGCTTACT**GA**
GCCCTCTGCTGCCACCACTTGTGACTGTACCCATGAGGTATGGAAGGAGCAGGACTG
GCCTGAGCATGAGCTGGAGAGTGTCTTGTCTAGCAGCTGGTTGGGACTATATTCTG
TCACTGGAGTTTGAATGAGCCACCCGCACTTGGTGTGCATGGGACATCTA
CTGGTCTGGGAAGCCACCCACCCAGGGCAATGCTGCTGTGATGTGCCCTTCCCTGAGTCC
TTCCATGTGGAGCAGAGGTGTGAAGGAATTTCAGTGGTTGTGATGCCAAATCACAGAAC
AGAATTTCATAGGCCAGGCTGCCGTGTTGACTCAGAAGGCCCTTCTACTTCAGTTG
AATCCACAAAGAATTAAAACCTGGTAACCCACAGGCTTCTGACCATTCATTGTTGGGTT
TTGCAATTGACCCAACCCCTGCTCACCTGAGGAGCTTCTTGGAAACCCAGGATGAAACT
TCCTCCCTGCCCTAACCTTCTACTCACCTATTGCTCTCTGTGATGTGCAACCTGAGCTG
GGAAAGGATTGGATGCCCTCTGTGGGCCCTGGGGCTGAGAACACACCTGCCATTAC
TGGCTTCATTAGGTGGCCCTAGGGAGATGGCTTCTGCTTGGATCACTGTTCCCTAGCAT
GGGCTTGGCTATTGGCATGTCATGCCATGCCCTCCCAATAGTCTTCAGGCCCTCAGTG
AAGTTTGGCTAAAGGTTGGTGTAAAAATCAAGAGAACGCTGGAGACATCATGGATGCCATG
GATTAGCTGCAACTGACAGCTCAGGTTGATCAAACAAAAGGAAACATTGTCATGTG
GTCTGACCATGTGGAGATTTCTGGACTTGTCTAGAGCTGCTTAGCTGATGTTTTGAGT
TACGATTTTGGAAATCCCACCTTGAGTCTGAAAGTGTAGGAAGGCTTCTTCAACCTT
GGGCTTGGATATTGCCACAGAAGAATTGGCTTTTTCTTAATGACAAGAGACAGT
TGCTGTTCTCATGTTCAAGTCTGAGAGCAACAGACCTCATCATCTGTGCCCTGGAAAGAGTT
CACTGTCATTGAGCAGCACAGCCTGACTGCTGGCCTCTGTCACCCCTTATTCCACTGCC
TTTGACAAGGGTTACATGCTGCTCACCTACTGCCCTGGGATTAAATCAGTTACAGGCCAG
AGTCTCCTGGAGGGCTGGAACTCTGAGTCTCTCATGAAACCTGTAGCTGACCTAAAGAAAT
TCTTAAATCACCAGATGGAAACAAAAAAAGGGCGCCGCGACTCTAGAGTCG
ACCTGCACTAGGGATAACAGGGTAATAAGCTGGCCGCATGG

FIGURE 153

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911
><subunit 1 of 1, 348 aa, 1 stop
><MW: 39711, pI: 8.70, NX(S/T): 1
MAATLGPLGSWQQWRRCLSLARDGSRMLLLLLLLGSGQGPQQVGAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWEQVHFKIHGQGKKN
LHGDSLAIWYTKDRMQPQPGPFGNMDKFVGLGVFVDTYPNEEKQQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRNLYHDTFLVIRYVKRHTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKLFELETVERTPEEEKLHRDVFLPSVDNMKLPEMTAPLPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQRKRFY
```

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

09378544-101501

FIGURE 154

CCGAGCCGGCGCAGCGACGGAGCTGGGGCCGGCTGGGACCATGGGTGAGTGCAATCTACGGATCAGTCT
CTGATGGTGGCTGTTAACCTCACTGGGACTCCAAAGATTCCATGAAGAAAATCAGTGTCTCAITCAAGAT
TGGGGCTGCTCAGAAATCTGGCAGCTGGAAATCTGTGTTCTAGAAGAGGTTTAATTAAATGCCCTGAGTCT
GACATGTTCCGATTGGAGGTGAAACCATGAAGAGAAAATAGAATATTAAATATGCTTTCGCCACCGTCT
TGCTGCTGCTGCCCTGCCCTGCTGGCTTGTGAGGCTCACGCTGAGTCTTCACCTGATCCGGTGT
CGAGACCCGTTATGAAGCTTTTGACTGCACATCCCCAGTGTGCCCAGGAGCACAGGTAACCCACTGTATGTCATTC
CGCATCATTTAAAGCTGCTCAGTGATGTGTTCTGGCAGGAGCACAGGTAACCCACTGTATGTCATTC
AAACAAAGCGACCRGARATTGACATCTGGCTAACAGGAAACGGTATCACCAAACACTGGAAGCTTCA
TTAGTCACATGTCAAAAGATCGGAGCCCTTGTGAAAGGCTTGTGAACTCTTGTGTCAGGAACTTCA
CATTGTGAGATGAGACTGAGACGAGACGAGGAGTGTGAGCATTGAGAAGGGTCACTGCTGAGGATA
TCATATCAAAAGAACACAACCTCTGCCCATGATTGCTGTCAGCACAGGCTCATTTAGAGACCACGGAAAA
GCCGGACCCCTACAAGTGCGCTTGTGCTTATGGCTTCTCCAGAACAGGATTTGAGCTGGAAAGAAGATTATTTCA
GGCACCGCCAAGTGCGCTTGTGCTGTCAGGAGCTGCTATGGCTGCGTAAAGAACAGTATCTGGAAAGGAGC
AGGGCTGCTAGACTCTCTAGTGGAAACAGCAGGAGAGACCTACGGGAGATGCCAAGATCGTGG
ATGTCCTCCAAAGAACAGCTTGTGACTCTCCACCCATAGACTCTGCTCCACCTTGTGAAATGTC
TTCTCTGACCAAAATGGCTGTGATGACATGGGACTTCAAGGTTAAAGACCATCAGATCGGGATGAAA
GGGAAAGACGGAGAAATTTACTCGGGTATCTCTCTGGTGCACCCACCCATCTGAACCAAACCCATCG
GCCGGTGCAGCGTGCACCGGGAGGAGAGCTTCTCTCTACTCTGCTCATGATCTCAGTCTGACTCTGT
CACCAAGTTCTAGTGCCTTGGGCTTCAAGGAGTTCCAGGTTGAGCAGGTTGATCTTGTGAGCTT
GGCAAGACAGGAAAGGGAGCTGAATTCCTCCGGGACTTCTACATGGCTGAGTGTACACCTTCCACCC
CTTCTCTGCAAGACCCAAAGGGTCTCTCCAAAGGCCCCTGTGCCCTTGTGAAA
GGGACATGTGTTAGTGGTGGCTGAGCTGTCACAAATTATATTATGATGAGCTGTCAGGAAAGGATTCTAAAAGG
TATGAGCTACAGCATAGAATCCATGACAGCATAGGGAAAGGTCACCTTAGTTGTCTGTGTTAC
TARGGGTAGAGATTATGCTTTAAAGGCTAAATATTGTTGTGGGAACACAGATGGTTGGGTTGAGCAGT
AAAGCACATTGCTCAATGTGAGCTGTAATTGGCTGGTAACTTGGCTGGTAACTGGCAGTTCACAGGAGTAGAAGGACT
TATCATGGCCAGACTTCTGCTTAAAGGCAAGAATATATAGTCTGACACTGAGCTGAACTTGTGAC
TCTCTGCGCTTGTGCTTAAAGGTTCTGAGTGTATTCTAAATTAAGGTTGCCAAACTTGTCTGAAAG
GCCAGATTGTAATATTCAAGCTGTGTGAGCCTAACAGGCCCACATACAGTCTGTCTGAC
TTCTGAGCAGGAAAGGCCACCAAGACAGTACATAAAGGAAATATGTTGTAGCTGGGTTCCAGGGCAGACAAAACA
GATGGTGGACAGACTTGGGCTCTGGCTGTAGTTGTGACCTCTCTCTGAC
ACTTCCAGCACATTGAGAAGCAGTGAATACCAAGAAATTATTCAGTGGTCTCTCCAGTAACTTCTGCTAGAAACA
CAGAATTGGCTGTGATCTGACACTGAGAACAAAATCTGAGGTTAAATTAACATTGAGATTAGAATCATGAA
AACTGATGAGAAGATTACTTGATGTTATGATGATGTTGAGCTACAGATAGTTTAAGTATGTTCTAAATTGTT
CTGCTGAGCTGCTTATGCTGTATGCTGAAATTGTTGAGCTTGTGTTAGTTTATAGTTAGGAAAGTATT
TCTCTAAGGACAGTTTAGATGACTCTTATCTGCTGAGTAAATTCAATTGCTGACCTGCTGGTGGTTAGAAG
GAGGCTGAGAAGATGAATTAGGCACTTCTCCAAATAAAACCTAATTAGGCTCATCCCTTGCACAGCTGAGA
ACTGGATGATTCTTAAACCATTTCTCATGCTTCAATGTTAAATTGAGATTGTTAAATGCGTTTGG
AGAACATTGCTATTAGGAGTTACAGATCTTATAAGGTGTTTATATTAGAAGCAATTATAATTACATCTG
TGATTCTGAACTAATGGCTAATTGAGAAATGGAAGTGAAGTGGAGATTCTGTGTCATGGCATTCC
AACTTTTCTCTTGTGTTGGCTGAGTGTGATTTCTGATTAATAAAATTGAAAGATA

FIGURE 155

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329
><subunit 1 of 1, 480 aa, 1 stop
><MW: 55240, pI: 9.30, NX(S/T): 2
MLFRNRFLLLALAALLAFVSSLQFFHLIPVSTPKNGMSSKSRKIRMPDPVTPEPVTDPPVY
BALLYCNIPSVAERSMEGHAPHFKLVSVHFIRHGDRYPLVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKGSGASFESPILNSLPLYPNHPLCEMGELTQTGVVQHLQNGQLLRDIYLK
KHKLLPNWDWSADQLYLETTGKSRTLQSGLALLYGFLPDFDWKKIYFRHQPSALFCSGSCYCP
VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVTPKQLRAANPIDSMLCHFCNVSFPC
TRNGCVDMEHFVKVIKTHQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELF
ALYSAHDVTLSPLSALGLSEARFPRFALARIFELWQDREKPSHEHSVRILYNGVDVTFHTSF
CQDHHKRSKPKMCPLENLVRFVKRDMFVALGGSGTNYYDACHREGF
```

Signal sequence:
amino acids 1-18

FIGURE 156

FIGURE 157

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306
><subunit 1 of 1, 916 aa, 1 stop
><MW: 100204, pI: 4.92, NX(S/T): 4
MIPARLHRDYKGLVLLGILLGTLWETGTCIRYSVPPELEKGSRVGDISRDLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDIILMEDKVKIYGVEVEVR
DINDNAPYFRESELEIKISENAATEMRFPLPHAWDPD1GKNSLQSYELSPNTHFSLIVQNAGA
DGSKYPELVLRKRALDREEKAHHHLVLTASDGGDPVRTGTARIVMVILDANDNAPAFQAQPEYR
ASVPENLALGTQLLVNNATDPDEGVNAEVRYSFYVDDKAQQVFKLDCNSGTTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSIASSVPENS PRGTLIALLNVN
DQDSEENGQVICFIQGNLPFKLEKS YGNYYSLVTDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLNVA DTDNPPVFPQASY SAYIPEENNPRGVSLVSVTAHDPDCEENAQITYSLAENTI
QGASLSSYVSINS DTGVLYALSSFDYEQFRDLQVKVMARDNGHPPLSSNVSLSLFVLDQNDN
APEIYLPA LPTDGSTGVELAPRSAEPGYLVTKVVAVD RD SGQNAWL SYRLLKASEPGLFSVG
LHTGEV R T AR ALL DR AL K QSL VV AV QDH GQPPLS AT VT LT V A VAD SI P QV LAD LG SLES PA
N SET SD L TL YLV VAVA AVSCVFL AFV ILL AL RL RR WH KS R LL Q AS GGG LT G A PASH FVGVD
GVQ AFL Q TY S HEV S LTT DS R K SH L I FP Q P N Y ADML V S Q E S F E K S E P L L S G D S V F S K D S H G L
IEV S LY Q I F F L F F N C S V S Q A G V Q R Y D H S S L R P Q T P R L K Q L S H L C L R C N R D Y R C K P P T V C L S
I Y L S I Y L S I Y L L S C T D G S L T P V I P V L W E A E A G G S P E V G S L R P A
```

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

FIGURE 158

CCCCAGGCTCTAGTGCAGGGAGAAGGAGGAGGCAGGAGGTGGAGATTCCCAAGTTAAAG
GCTCCAGAACATCGTGTACCAGGCAGAGAACACTGAAGTACTGGGGCCTCCACTGGGCCGAA
TCAGTAGGTGACCCGCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCTGCTCTGCTGGGGGAGCCTGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGTGATGAGTGCCAACCCCATTCGAGCCTGGCAGGCCG
CCTTGTTCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTGTAGGTGGCAACTGGGTCTT
ACAGCTGCCCACTGTAAAAACCGAAATACACAGTACGCCCTGGAGGACACAGCCTACAGAA
TAAAGATGGCCCAGAGCAAGAAATACCTGTGGTTCACTGCATCCCACACCCCTGCTACAA
GCAGCGATGTGGAGGACCAACCATGATCTGATGCTTCATACTGCGTGACCAGGCATCC
CTGGGGTCAAAGTGAAGGCCATCAGCCTGGCAGATCATTGACACCAGCCTGGCAGAAGTG
CACCGTCTCAGGCTGGGGCACTGTACCCAGTCCCCGAGAGAAATTTCCTGACACTCTCAACT
GTGCAGAAGTAAAATCTTCCCCAGAAGAAGTGTGAGGATGCTTACCGGGGCAAGATCACA
GATGGCATGGCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCAGATTCTGGAGG
CCCCCTGGTGTGATGGTCACTCCAGGGCATCACATCCTGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGCTGATTCTAGGATAAGCACTAGATCTCCCTTAATAAAACTCACAACTCT

CTGGTTTC

09078441.101601

FIGURE 159

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336
<subunit 1 of 1, 260 aa, 1 stop
<MW: 28048, pI: 7.87, NX(S/T): 1
MGRPRPRAAKTWMFLLLGGAWAGHSRAQEDKVLGGHECQPHSQWPWQAALFQQGQLLCGGVL
VGGNWVLTAHCKPKYTVRLGDHSLQNKGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDSGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRY
LDWIKKIIGSKG
```

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

FIGURE 160

GGCGCCGGTGCACCGGGCGGGCTGAGCGCCTCTGCGGCCGGCCTGCGGCCCGGGCG
CGCGCCGCCAACCCCCAACCCCAGGCTAGCGCTCCGCCGCCGAGGCCCGGCCGG
CCCCGGCCGCCGGGGAAACCGGGCGATTCTCTGCCGCGCTCAAACACCTGATCCCATAAAAC
ATTCACTCTCCGGGCCGGCCGGCTGCGCACCCGCGGCCAGGCCAGCGAGGCCAG
GGAGCGCGCCGAGCCCTCGCTCCCGCGGCCGGGCCGGCGTAGCGGCCGCGCTGGA
TGCGGACCCGGCCGCGGGAGACGGGAGACGGGAGGCCGGCGTAGCGGCCGCGCTGGA
CCCCCCAACCCCTACGTGAAGAGGGCGTCCGCTGGAGGGAGGCCGGCTGCTGGCATGGGTG
CTGTGGCTGCAAGGCTGGCAGGTGGCAGCCCATGCCAGGTGCCCTCGTATGTCACAATGA
GCCAAGGTGACGACAAGCTGCCAGCAGGGCTGCAAGCTGTGCCGTGGCATCCCTG
CTGCCAGCCAGCGCATCTCTGCCAGGCCAACCGCATCTGCCAGCTGCCAGCTTC
CGTGCCTGCCAACCTCACCATCTGTGGCTGACTCGAACATGTGTCTGCCCGAATGATGC
GGCTGCCTTCACTGGCTGGCCCTCTGGAGCAGCTGGACCTCAGCATAATGCACAGCTCC
GGTCTGTGGAACCTGCCACATTCACGGCTGGGCCCTACACCCCTGCACTGGACCGC
TGCGGCTGCAAGAGCTGGGCCGGCTGGCTGGCCGCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACACCGCGCTGAGGCACTGCCATGACACCTTCCGCCACCTGGCAACCTCA
CACACCTCTCTGCAACGCCACCGCATCTCCAGCGTCCGCCAGGCCCTCCGGCTG
CACAGCCTGACCCGCTCTCACTGGCACCAGAACCGCTGGCCCATGTGCACCCGATGCCCT
CCGTGACCTTGGCCCTCTGACACCTCTATCTTTGCCAACATCTATCAGGCCCTGCCA
CTGAGGCCCTGGCCCCCTGCGTGCCTGCACTGAGGCTCAACGACAACCCCTGGGTG
TGTGACTGCCGGCACGCCACTCTGGGCTGCTGCAAGAGTTCCCGGGCTCTCCCGA
ACCTGCCCCCTGCAACCGCTGGCCACCGGCTGGCTGACCTCAAACCCCTAGCTGCCAAATG
ACCTGCAAGGGCTGCGCTGTGGCCACCGGCTTACCATCCCATCTGCCAGGCCACCC
GATGAGGAGGCCGCTGGGCTTCCAAAGTGTGCTGCCAGCAGATGCCGTGACAAGGCTCAGT
ACTGGAGGCCGAGAACCCAGCTGGCAGGCCAATGCGTGAAGGGAGCGCTGCCGGGTG
ACAGCCGCCGGCAACCGCTCTGGCCACGCCACATCAATGACTCACCTTTGGACTCTG
CCTGGCTCTGCTGAGCCCCGCTCACTGCACTGCCAGGGCTCCGAGCCACAGGGTT
CCCCACCTCGGGCCCTCGCCGGAGGCCAGGCTTCAACCGCAAGAACCCGACCCGCA
GCCGTCTGGCCAGGAGCGCAGCGGGGTGCGGGACTGACTCGAACAGGCTCAGGTGCC
CTACCCAGCCCTCACCTGCAAGCTCACCCCCCTGGGCTGGCGTGGTGTGACAGTGT
TGGGCCCTGCTGCACCCCCAGGGACACAAGAGCGCTGCAAGCAGGTGTGTACATAC
GGGGTCTCTCCACGCCCAAGCCAGCCGGGGCCGACCCCTGGGGCAGGCCAGGCCAG
GTCTCTCTGATGGACCCCTGCCGCCACCCCCATCTCCACCCCCATCATGTTACAGGG
TTCGGCGGCGAGCTTGTGAGAACGCCGCTCCACCCAGATCGCGGTATATAGAGATAT
GCATTTTATTTACTTGTAAAAATATCGGACGACGTGAAATAAGAGCTTTCTTAA
AAAA

FIGURE 161

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184
><subunit 1 of 1, 473 aa, 1 stop
><MW: 50708, pI: 9.28, NX(S/T): 6
MKRASAGGSRLLAWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPGVIPAASQRI
FLHGNRISHVPAASFRACRNLTILWLHSNVLARIDAAAFTGLALLEQLDLSDNAQLRSVDPA
TFHGLGRHLTLHLDRCGLQELGPGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLLLHQNRVAHVPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRLNDNPWVCDRCARPLWAWLQKFRGSSSEVPCLPQQLAGRDLKRLAANDLQGCA
VATGPYHPIWTGRATDEEPLGLPKCCQPDADKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAPPPLTAVRPPEGSEPPGFPTSGPRRRGCSRKNRTRSHCRLGQA
GSGGGGTGDSEGSGALPSLTCSLTPLGLALVLWTVLGPC
```

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

FIGURE 162

GGAAAGTCCACGGGGAGCTTGGATGCCAAGGGAGGACGGCTGGGTCTCTGGAGAGGGACTAC
TCACTGGCATATTCTGAGGTATCTGAGATAAACACAGCCTCAGATACTGGGGACTTTAC
AGTCCCCACAGAACCGTCTCCCAGGAAGCTGAATCCAGCAAGAACATGGAGGGCAGCGGGG
AGCTCATTTCGACAGAACAGGCAAGTCTTTCTCTCTCTCTCTGGCTTATCTCTGGCG
GGCCCGGGCGAACCTAGAAGCTATTCTGGTGGAGGAACCTGAGGGCAGCTCCTTGTAC
CAATTAGCAGAACAGGACCTGGCTCTGGAGCAGAGGGAAATTCTCAGGGGGGGTTAGGGTTG
TTTCCAGAGGGAAACAAACTACATTGAGCTCAATCAGGAGACCGCGGATTGTTGCTAAAT
GAGAAATTGGACCGTGGAGATCTGTCGGTCAACAGAGGACCTGTGCTACAGTTTCAAGT
GTTGCTAGAGAGTCCCTTCAGAGTTTCTCAAGCTGAGCTGCAAGTAATAGACATAAACGACC
ACTCTCAGATTCTGAGAACAAATGTTGGTGAAGATCTCAGAGAGCAGTCTCCTGGG
ACTACGTTCTCTGAAGAATGCCGAAGACTTAGATGTAGGCCAAAACAATATTGAGAACTA
TATAATCAGCCCCAACCTCTTCTGGCTCTCACCGCAAACAGCAGCTGATGGCAGGAAAT
ACCCAGAGCTGGTCTGGACAAAGCGCTGGACCCAGAGGGAAAGCTGAGCTCAGGTTAAC
CTCACAGCACTGGATGGTGCTCTCCGCCAGATCTGGCACTGCTCAGGTTACATCGAAGT
CCTGGATGTCAGATAATGCCCTGAATTAGCAGCCTTTCTATAAGTGTGAGACTCTG
AGGACAGTCGGTGAAGGTTCTGGTGTGAAGGTTCTGCCACGGATGTAGACACAGGACT
AACGGAGAGATTCTCAATTCTCAAGCTCAGAGAGATTGCAAACACCTTTAAAGAT
CAATCCTTGACAGGAGAAATTGAACATAAAAAAAACAACCTGATTTCGAAAAAAACTTCAGTCT
ATGAAGTCAATTGAGGCAAGAGATCTGGAACTTTCTGGAAAATGCACTGGTCTGATT
CAAGTGTAGATGTGAACCATGCCCCAGAATTACCATGCTGCAATTACAGGCCAAT
ACCTGAGAACGCCCTGAACCTGTGGTTGCACTTTCACTGTTCAAGATCTTGATTAGGAG
AAAATGGGAAAATTAGTGTCTCATTAGGAGGATCACCTTCTCTGAAATCCGGGAA
AACTTTTACACCTACTAACGGAGAGACACTAGACAGAGAACAGCAGGAAATAACACAT
CCTACTCACTGCACTGACTTGGGGACCCCTATGCTGATAACACAGCTAACATGACCGTGC
TGATGCCCGATGTCATGACAACGCTCCGCCCTTCACCAAAACCTCTACACCCCTGTCGTC
CGCGAGAACACAGGCCCTGCACATCGCAGCGTCAAGCAGAGACTCAGG
CACCAAGCCCAAGGTCACTTACTCGCTGCTGCCCTCAGGACCCGCACCTGCCCTCACAT
CCCTGGTCTCCATCAACCGGGACAACGCCACCTGTCGCCCTCAGGTCTCTGGACTACGAG
GCCCTGCAGGGGTTCCAGTCCGGCTGGCGCTTCAGACACCAGGGCTCCCGGGCTGAGCAG
CGAGGCCCTACCTCCGCTCTCCGGGAGGGCCCGACCCAGGGCCAGGCGACTTGCCTCACCG
TCTACCTGGTGGTGGCGTGTGGCGACTCCGGCCAAGACGCCCTGTCGCT
GCTGTCAGAACGGCAAGGTGGTGGCGACTCCGGCCAAGACGCCCTGTCGCT
GCTGTCAGAACGGCAAGGTGGTGGCGACTCCGGCCAAGACGCCCTGTCGCT
CCGGCAGGCTCTGAGCGAGCGAGCGGCAACGGCAGACAGCTGGTGTGCTGGTCAAGGAC
AATGGCGAGGCTCCGCGCTCGCCACGCCACGCTGACGTGCTCTGGTGGACGGCTCTC
CCGGCCCTACCTCCGCTCTCCGGGAGGGCCCGACCCAGGGCCAGGCGACTTGCCTCACCG
TCTACCTGGTGGTGGCGTGTGGCTCTGGTGTCTTCTCTTCTGGTGTCTCTGGT
GTGGCGGTGCGGTGTGTAGGAGGAGCAGGGCGGCCCTGGTGGGTGCTGCTTGGTGGCGA
GGGGCCCTCCAGGGCATCTTGTGGACATGAGCGCACCAGGACCTATCCAGAGCTTAC
AGTATGAGGTGTGTGGAGGAGGCTCAGGGACCAATGAGTTCAAGTTCCTGAGCCGATT
ATCCCCAACCTCCCTCCCATGCCCTGGGAAAGAAATTAACAGGAAATTCTACCTTCCCAA
TAACCTGGGTTCAATATTCACTGCAAGTTTATTCCCTTAAAGTGTAAATTGTAACGGAT
TTACTCTGATTCTCATGTTCTCCCTTGTGTTAAAGTGAACATTACCTTATT
CCTGGTTCTT

FIGURE 163

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314
<subunit 1 of 1, 798 aa, 1 stop
<MW: 87552, pI: 4.84, NX(S/T): 5
MEASGKLICRQRQVLFSLLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSRRGNKLHLQLNQETADLLLNEKLDREDLCGHTEPCVLRFQVLLESPFEFFQAELQV
IDINDHSPVFLDKQMLVKVSESSPPGTTFPLKNAEDLDVGQNNIENYYIISPNSYFRRVLTRKR
SDGRKYPELVLKDALKDREEAEELRLTLTALDGGSPPRSGTAQVYIEVLDVNDAPEFEQPFY
RVQISEDSPVGFLVVVKVSATDVTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSPIPENAPETVVALFSVS
DLDSEGENGKISCSIQEDLPFLLKSAENFYTLTERPLDRESRAEYNITITVTDLGTPMLITQ
LNMTVLIADVNDNAPAFQTSYTTLFVRENNSPALHIRSVSATDRDGTNAQVTYSILLPPQDP
HLPPLTSVLVSINAQNQHGLFALARSLDYEAQGFQFRVGASDHGSPALSSEALVRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELGLFGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAQ
ADLLTVYLVALASVSSLFPSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMMSGTRT
LSQSQQYEVCLAGGSGTNEFKFLKPIIPNFPPQCPGKEIQGNSTFPNNFGFNIQ
```

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site,

amino acids 418-421, 436-439, 567-570 and 786-789

FIGURE 164

ACCCACCGTCCGCCAACCGTCCGCCAACCGTCCGCCAACCGTCCGCCGTAGCCGTGC
GCCGATTGCCCTCGGCCGGCAATGGTCCCGGCTGCCGGTGCAGGACCGCCCCGGTCAT
GGCGCTCCTCGGCTGGTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCGTCCGCCGTGG
AGGTTGAGAGGAAAGTGGCCTTATGGTCAAGAGGAGCAGCCTGCTCACCTCTCCAGGTG
GGGGCTGTGTAACCTGGGTGAGGAGGAGCTCTGCATGACCCGATGGCCAGGACAGGGCAGC
AGAAGAGGCCAATGCCGGTCTGGGCTGGACACCCAAGGGCATCACATGGTGTGCTG
TGATTCTGGGAAGCTGAGGACAAGTGGATTCAAGCCTAGCGGCGTCACCTGTGGTG
GGAGGAGCGGAGGACTCAAGTGCAACGTCCGAGAGAGCCTTCTCTGGATGGCGTGG
AGCACACCTCCCTGACAGAGAAGAGGAGTATTACAGAGCCAGAGTGGCGGAATCTGACG
CAGCCCCGACAGAGGACTCAATAAACACTGAAAGTCTGAAATCCCCAAGGTGAACGTGAG
GAGAGAAACATTACAGGATTAGAAAATTCACTCTGAAAATTAAATATGTCACAGGACCT
TATGGATTTCTGAACCCAAACGGTAGTGA~~CT~~TACTCTAGTCTGTTTACACCCGTGGT
GCCGCTTTCTGCCAGTTGGCCCTCACTTAACTCTGCCCGGCATTCCAGCTCTT
CACTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTCTACCAGGTTGGCACCGTAGC
TGTTCTAATATTTATTCAAGGGCTAAACCAATGCCAGATTAAATCATACAGATC
GAACACTGGAACACTGAAAATCTCATTAAATCAGACAGGTATAGAACCCAAGAAGAAT
GTGGTGGTA~~ACT~~CAAGCCGACCAAAATAGGCCCTCTCCAGCACTTGATAAAAAGTGTGGA
CTGGTTGTTGTATTCCATTCTTTAAATTAGTTTATTATGCTACCTTCA
CTGAGAGTATTGGTGGCTAATTCCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT
GAAAGAAGTGGAAAGAGGAACCTCAATCCTCGTTCAAGAAATTAGTGC~~T~~ACAGTTCTATA
CATTTCTCCAGTGACGTGTTGACTGAAACTTCAGGGCAGATTAAAAGAATCATTGTTGAA
CAA~~T~~GAATGTATAAAAAAATTATAAAACTGGTGT~~TTA~~ACTAGTATTGCAATAAGCAAATGC
AAAAAATTCAATAG

FIGURE 165

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333
><subunit 1 of 1, 360 aa, 1 stop
><MW: 39885, pI: 4.79, NX(S/T): 7
MVPAAAGRPPRVMRLLGWWQVLLWVLGLPVVRGVVAEESGRLWSEEQPAHPLQVGAVYLGEE
ELLHDPMGQDRAAEEANAVLGLDTQGDHMVMLSVIPGEAEDKVSSEPSGVTCGAGGAEDSRC
NVRESLFLSLDGAGAHFPDRREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCERNITGLE
NFTLKLILNMSQDLMDFLNPNPGSDCTLVLYTTPWCRFSASLAPHFNSLPRAFPALHFALDAS
QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTELETLKIFIFNQTGIEAKKNVVVTQADQ
IGPLPSTLIKSVDWLLVFSLFFLISFIMYATIRTESIRWLIPGQEQQEHVE
```

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to dilsufide isomerase

amino acids 212-302

N-glycosylation site.

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296

Thioredoxin domain

amino acids 211-227

FIGURE 166

CCCGGCTCCCTCCCTCGCCCCCTGGGGTCGCGCAGCAGATGCTGCAGGCCCTGGCT
CGCTGCTGCTGCTCTTCCCTCGCCTCGACTGCTGCCTGGCTCGGCCGCGGGCTCTCCCTC
TTGGCCAGCCGACTTCTCTACAAGCGCAGCAATTGCAAGCCCACCCGGTCAACCTGCA
GCTGTGCCACGGCATCGAACATACCGAGAACATGCGCTGCCAACCTGCTGGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGGCGCTGGATCCCGCTGGTCAATGAAGCAGTGCCACCCG
GACACCAAGAAGTTCTGTGCTCGCTCTCGCCCCCGTCTGCCCTCGATGACCTAGACGAGAC
CATCCAGGCCATGCCACTCGCTCTCGTGAGGTGAAGGACCGCTGCCCGGGTCAATGTCGG
CCTTCGGCTTCCCTGGCCGACATGCTTGAGTGCAGCTTCCCCCAGGACAACGACCTT
TGCATCCCCCTGCTAGCAGCAGCACCTCCCTGCCAGCCACCGAGGAAGCTCAAAGGTATG
TGAAGCCTGAAAAATAAAATGATGATGACAACGACATAATGAAACGCTTGTAAAAATG
ATTTGCACTGAAAATAAAAGTAGAGGAGATAACCTACATCAACCGAGATAACCAAAATCATC
CTGGAGACCAAGAGCAAGACCATTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAA
ATCGGTGCTGGCTCAAAGACAGCTTGCACTGACCTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGTCAATGGACAGAAAACAGGGTGGGAGCTGGTGAATCACCTCGGTGAAGCGG
TGGCAGAAGGGCAGAGAGAGTTCAAGCGCATCTCCCGAGCATTCCAGCTGCAGTGCTGA
GTCCCGGCATCTGATGGCTCGACAGGCCGCTCCAGAGCACGGCTGACCATTCTGCTCC
GGGATCTCAGCTCCCGTCCCAAGCACACTCTAGCTGCTCCAGTCTCAGCCTGGCAGCT
TCCCCCTGCCTTTGCACGTTGCATCCCCAGCATTTCTGAGTTATAAGGCCACAGGAGTG
GATAGCTTTTACCTAAAGAAAAGCCCACCGAATCTGTAGAAATATTCAAACATAATA
AAATCATGAATATTTAA

FIGURE 167

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920
><subunit 1 of 1, 295 aa, 1 stop
><MW: 33518, pI: 7.74, NX(S/T): 0
MLQPGPSLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMQKQCPDTKKFLCSLFAPVCVLDDLDTEIQPCHSLCVQVKDR
CAPVMSAFGPWPDMLECDRFPQDNDLCIPLASSDHILLPATEEAPKVCCEACKKNNDNDIM
ETLCKNDFAALKIVKEITYINRDTKIILETKSHTIYKLNGVSERDLKKSVLWLKDSDLQCTCE
EMNDINAPYLVGMOKOGGLELVIITSVKRWOKGOREFKIRSRSIRKLOC
```

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus
amino acids 6-153

FIGURE 168

GTGGAGGCCGCCGACGGATGGCGGGGCCGACGGAGGCCAGACGGGGTTGGCCGAGCCCCGGG
CCCTGTGCGCGCACGGGCCACCGCACCTACGCGCCGCTGGTGTTCCTGCTCGCGATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTGACCTGTGGCTGACGTAT
TGCTGAGGACTTGGTCCTGTCATGGAGCAGATCAACTGGCTGTCACTGGCTACCTCGTGG
TATCCACCCCATTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTCGGCTCCGTGCGGCC
ACCATCCTGGGTGCGTGGCTGAACATTGGCGGGAGTTGTGCTACGCATGGTGCCCTGCATGGT
TGTTGGGACCCAAAACCCATTGCTTCTCATGGGGCCAGAGCCTGTGCCCCTTGCC
AGAGCCTGGTCATCTCTCCAGCCAAGCTGGCTGCCCTGGTTCCAGAGCACAGCGA
GCCACGGCCAACATGCTGCCACCATGTCGAACCCCTGGCGTCTTGCGCCAAATGTGCT
GTCCCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCCGTTAATGCTCGGTGTCTATACCATCC
CTGCTGGCGTCTGCTGCCCTGTCACCATCTGCCCTGTGGAGAGTTGTGCCCTAACCCCG
CCCTCTGCCGGGCTGCCAGCTCCACCTCAGAGAAGTCTGGATGGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCATGTCATCTGGCTGTGCTTGGGGGAATGATCGGGATCTCTG
CCAGCTTCTCAGCCCTCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGTTTCC
GGCCCTCTGTGGCCTCTCTCATCACGTTGGGATCTGGGGCAGCTGGCTCTGGCCCCCTA
TGTGGACCGGACCAAGCATTCACTGAGGCCACCAAGATTGGCCTGTGCCCTGTTCTCTCTGG
CCTGCGTGCCTTGGCTGTCCAGCTGCAGGGACAGACCCCTTGCCCTGGCTGCCACC
TGCTCGCTGCTGGGCTGTTGGCTTCTCGGTTGGCCCCCTGGCCATGGAGTTGGCGGTGCA
GTGTTCTTCCCGTGGGGAGGGGGCTGCCACAGGCATGATCTTGCTGGGGCAGGCCG
AGGGAATACTCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGGAGCCGTCCTG
TCCACCTGCCAGCAGGGGGAGGATCCACTTGACTGGACAGTGTCTCTGCTGCTGATGGCCGG
CCTGTCACCTTCTCAGCTGCATCTGGCGTCTTCTTCCACACCCCATACGGCGCCTGC
AGGCCGAGCTGGGGAGCCCCCTCCACCGTAACGCCGTGGCGGCCAGACTCAGGGCCG
GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGCTCTGGGCCAGCACGGCAGCTCCGGA
GTGCAAGGCCAGGGGGCCCTCGCTAGAGGACCCAGAGGGCCGGAGCCCCCACCGCC
GCCACCGAGGCACTCCCCGTGCCAGGCCAGCAGGCCACCGACGCCCTCCGCCCGG
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTATTGACCCGGCTGGGCTCACTCCCTT
CTCCCTCCCGTGGGTGATCACGTAGCTGAGCCCTTGAGTCCAGGTTGCCGCCACATCGA
TGGAGGCGAACTGGAACATCTGGTCCACCTGCGGGCGGGGCAAAGGGCTCCTGCGGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

FIGURE 169

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988
><subunit 1 of 1, 560 aa, 1 stop
><MW: 58427, pI: 6.86, NX(S/T): 2
MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSPFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFGVAAIWILDHSVGLRAATILGAWLNFAGGSVLRMVPCMVGQTQN
PFAFLMGGQSCLCALAQSLVIFSPAKLAALWFPEHQRTANMLATMSNP LGVLVANVLPVVLV
KKGEDIPLMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNK
YVILAVCLGGMIGISASFSALEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK
HFTEATKIGLCLFLSLACVPFFALVSQLQGTLALAATCSSLGLFGFSVGPVAMELAVECSFPV
GEGAATGMIFVLGQAEGILIMLMALMTALTVRRSEPSLSTCQQGEDPLDWTVSLLLMAGLCTFF
SCILAVFPFHTPYRRLQABSGEPPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG
ASLEDPRGPSPHPACHRATPRAQGPAAATDAPSRPGRILAGRQASRFIDPAGSHSSFPWVIT
```

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,
280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

FIGURE 170

FIGURE 171

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331
<subunit 1 of 1, 1184 aa, 1 stop
<MW: 129022, pI: 5.20, NX(S/T) : 5
MMQLLQLLLGPGGGYLFLLGDCQEVTTLTVKYQVSEEVPSGTIVIGKLSQELGREERRRQA
GAAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFVDLATGDLALIHVEIQ
VLDINDHQPRFPKGEGELEISESASLRTRIPLDRALDPDTGPNTLHTYTLSPSEHFALDVIV
GPDETKHAELIVVKELDREIHSFFDVLVTAYDNGNPPKSGTSLVKVNVLDSNDNSPAFAESS
LALEIQEDAAPGTLIQLKTATDPDQGPNGEVEFFLSKHMPEVLDTSFSDIAGTGQVILRRPL
DYEKNPAYEVDVQARDLGPBPNIAPAHCVKVLIKVLDVNDNIPSIHVTVASQPSLVSEALPKDSF
IALVMADDLDSGHNGHNLVHCWLSQLEGLHFRLKRTNGNTYMLLTNTLDREQWPKYTLTLAQD
QGLQPLSAKKQLS1Q1SDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLIGINGKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNYYEAMGFEFQVIAEDSGQPMЛАSSVSVWVSSLDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLPAGTDTPPLATHSSRPFLTTT
IVARDADSGANGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGS
PPLQTRALLRVMFVTSDHRLRSARKPGALSMSMLTVICLAVALLGIFGLLIALFMSICRTEK
KDNRAYNCRAEESTYRQQPKRPQKHIQKADITHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA
GWDPCLQAPFHLPPTLYRTLRNQGNQGAESREVLPQDTVNLLFNHPRQRNASRENLNLPPEP
QPATGQPRSRSLKVAGSPTGRLLAGDQGSEEAPRPPASSATLRRQRHLNGKVSPEKESGPRQ
ILRSLVRLSVAAFAAERNPVEELTVDSPPVQQISQLLSSLHQGQFQPKPNHRGNKYLAKPSS
RSAIPDTDGPSARAGGGQTDPQEEEGPLDPEEIDLSPVKQLLIEELSSLLDPSTGLALDRLSAPD
PAWMARLSLPLTTNYRDNVISPDAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSL
LEMMLLEQRSSMPVEAASEALRRLSVCGRTLSLTLATSAASGMKVQGDPGGKTGTEGKSRGSS
SSSRCL
```

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

FIGURE 172

CGGACGCGTGGGCGGACCGTGGGGAGAGCCGCAGTCGGCTGCAGCACCTGGAGAAGG
CAGACCGTGTGAGGGGGCTGTGGCCCCAGCGTGTGGCTCGGGAGTGGAAAGTGGAG
GCAGGAGCCTCCTTACACTTCGCCATGAGTTCTCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTTGAGTTGGGTGCTTCTCATGCCAATTGTTAAAGACTAT
GAGATACGTCAGTATGTTGACAGGTGATCTCTCCGTGACGTTGATTTCTGACCACAT
GTTGAGGCTCATCATCTTGAATCTTAGGAGTATTGAATAGCAGCTCCGTTATTTCACT
GGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTATGGTGCCTTTACATTGGC
TATTTTATTGTGAGCAATACCGACTACTGCATAAACACAACGACTGCTTTTCTGCTCTT
ATGGCTGACCTTATGTATTTCTGGAAACTAGGAGATCCCTTCCATTCTCAGCCCAA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCGGGTTGGTGTGATTGGAGTACTCTC
ATGGCTCTTCTTCTGGATTGGTGTGCAACTGCCATACACTTACATGTCCTACTCCCT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGGACTGCTGCAAACCATGGATA
TGATCATAAGAAAAAGAAAAGGATGGCAATGGCACGGAGAACATGTTCCAGAAGGGGAA
GTGCAATAACAAACCATCAGGTTCTGGGAATGATAAAAAGTGTACCTCAGCATCAGG
AAAGTGAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTGGAGAACATTAAGCAGGCAGC
TTTTCTGGAACAGCTGATCTATGCTACCAAGGAGAGAACATAGAACATCCAAAACCTTC
AAGGGAAATTTTAAATTCTGGTACTTTCTCTATTACTGTGTTGGAAATTTT
CATGGCTACCATCAATATTGTTTGATCGAGTGGAAAACGGATCTGTCAACAGAGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGTGAAGTTGGTCCACACATT
TCCTTCATTCTGGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC
CAAGTCTTTATGCCATCTAGCAGTAAGTCTCCATGTCATTGCTCTGCTATTAGCAC
AGATAATGGCATGTACTTTGTCCTCTGTGCTGATCCGAATGAGTATGCCATTAGAA
TACCGCACCATATAACTGAAGTCCTGGAGAACACTGCACTTCACCTCTATCACGGTGGTT
TGATGTGATCTCTGGTCAGCGCTCTCTAGCATACTCTCTCTATTGGCTCAAAAC
AGGCACCAAGAGAACAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCGAGT
GGTTCAAAATTAGATAAAGAGGGGGAAAATGGAACCAAGGGCTGACATTATAAAC
AAACAAAATGCTATGGTAGATTTCACCTTCATAGCATACTCCTCCCGTCAGGTGATA
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGAGAACTAACCTCAAGAACACTCA
GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGGAGCCAAGAACACTAA
AGGTGAAAATACACTGGAACCTGAGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGT
AGGATTCCGTTTAAGGTTACATGGAAAAGGTTATAGCTTGCCTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGAAGCTTGGCCCATGGCCAACTGTTTATTGCACTTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEI
LGVLNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFF
WKLGDPPPILSPKHGILSIEQLISRVGVIGVTLALLSGFGAVNCPTYMSYFLRNVTDTDI
LALERRLLQTMDMIISKKRMMARRTMFQKGEVHNKPSGFWMKSVTTASGSENLTLIQ
QEVDALLELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF
DRVGKTDPVTTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISS
SKSSNVIVLILLAQIMGMYFVSSVLLIRMSMPLYERTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTCCTTACACTGCCATGAGTTCTCATCGACTCCAGCA
TCATGATTACCTCCCNGANACTATTTTGATTTGGCTTCTTCNGCGCCAATGTT
TAAAGACTATGAGATACTGTCAGTATGTTGACNGGTGATCTCTCCGTGACGTTGCCATT
CTTGCACCATGTTGAGCTCATCTTGAAATCTTNGGAGTATTGAATAGCAGCTCCCGT
TATTTTCACTGGAAAATGAACCTGTGTGAATTCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACACGACTGCTTTTT
CCTGTCCTTATGGCTGACCTTATGTATTCCAG

FIGURE 175

GTTGGCCCTGGGGAGGGGAAGGGAGCCNGGCCCTTCCTAAATTTGGCAAGGGTTTC
TTTNTTGAAATCCGGGTTNNGNATACTTCCCAGAAAATTTTTGGATTTGGGTAGNTT
TTTTTCATCGCCAATTGTTAAAGACTATGAGATAACGTCAGTATGTTGACAGGTGATNTT
NTCCGTGACGTTGCATTTCTTGACCATGTTGAGCTCATCATNTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCCGTTATTTCACTGGAAAATGAACCTGTGTAAATTCTGCTGATC
CTGGTTTCATGGTGCCCTTTACATTGGCTATTTATTGTGAGCAATATCCGACTACTGCA
TAAACAAACGACTGCTTTTCTGTCTNTTATGGCTGACCTTATGTATTNTNTGGAAAN
TAGGAGATCCCTTCCCATTCTC

FIGURE 176

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737
>subunit 1 of 1, 1141 aa, 1 stop
>MW: 124671, pI: 5.82, NX(S/T): 5
MAGARSRDWPAGSGICVLFGLSLLVELLFSRAVAFNLDMGALRKEGEPGSLFGFSVALHRQL
QPRPQSLLVGAQALPGQQANRTGGFLFACPLSLEETDCYRVDIDQGADMQKESKENQWL
GVSVRSQGPQGGKIVTCAHRYEARQVDQILETRDMIGRCFVLSDQLAIRDELDGGEWKFCEG
RPQGHEQFGFCQQGTTAAAFSPDSHYLLFGAPGTYNWKGATARVELCAQGSADLAHDDGPYE
GGEKEQDPRLIPVPAWSYFGSIDSCKGLVRAEFLSFVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSFGFGSYLAVADLNSDGWPDLIVGAPAYYFERQEELGAVVYLNQGHHWAGI
SPLRLCGSPDSMFGISLAVLGLDNQDGFPDIAVGAPFDGDGKVFIYHGSSLGVVAKPSQVLE
GEAVGIKSFGSYLSGSLSMDGNNQYPDLLVGSILADTAVLFRARPLILHVSHEVSIAPRSIDLEQ
PNCAAGGHSVCDLFCFSYIAVPSSYSPVTDALYVLDADTDRRLRGQVPRVTFLSRNLEEPK
HQASGTVVWLKHQHDRVCGDMFQLQENVKDKLRAIVVTLSYSLQTPLRQQAPGQGLPPVAP
ILNAHQPSTQRQAIEHFLKQGCGEDKICQSNLQLVHARFCTRVSDETFQPLPMVDGTTALFA
LSGQPVIGLELMVTNLPSDPAQPQADGDDAHEAQLLVMLPDSLHYSGVRALDPAEKPLCLSN
ENASHVCECELGNPMKRGQAFTYFLILSTSGISIETTELEVELLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSVKYEVTVSMQGQSLRTLGS AFLNIM
WPHEIANGKLLYPMQVELEGQGPQKGKGLCSPRPNIILHLDVDSRDRRRRELEPPEQQEPGE
RQEPMSSWWPVSAAEKKNITLDCARTANCVVFSCLYSFDRAAVLHVWGRLWNSTFLEEY
SAVKSLEVIKRANITVKSSIKNLMLRADASTVLPVMVYLDPMAVVAEGVPWWVILLAVLAGLL
VLALLVLLWKMGGFKRAKHPEATVPQYHAVKI PREDRQQFKEEKTGTILRNNWGS PRREGP
DAHPILAADCQHPELGPDPGHPGPCTA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCCGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCAGCAGGAGCTGCCGACAGTCCTGGCT
CACAAACAAGATGCTCAAGGTGTCAGCGTACTGTGTGTGCAGCCCTGGTGCAGTC
GTCTCTCGCAGCTGCCGGCGGTGGCTGCAGCGGGGGCGGTGGACAGGGTAATTTC
TGGATGATAAAACATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC
AAATTCGAGACGAAGTAGAGGATGATTATTCGCACTGGAGTCAGGAAAACCTTCGA
TCAGGCTTCTAGATCCAGCTAAGGATCCATGCTAAAGATGAAATGTAGTCGCCATAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCAGTCAGTCACACACAGGATG
AAAGAACAGGGAGTAGACCATAGGCAGTGGAGGGTCCATAATTATCCACCTGCAAGCAGTG
CCCAGTGGCTATCCAGGCCCTGTTGTGGTTCAGATGGTCATACCTACTCTTTCACTGCA
AACTAGAATACTAGGCATGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCCTTCAGATAAGCCCACCACTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAAGGAAAGTGGAAACAGATTGCGGGACTGGTTCAAGGCCCTCATGAAAGTGGAA
GTCAAAACAAGAACAAAAACATTGCTGAGGCCCTGAGAGAACAGATTGATACCAACAGCATC
TTGCCAATTGCAAGGACTCACTGGCTGGATGTTAACAGACTTGATACAAACTATGACCT
GCTATTGGACCAAGTCAGAGCTCAGAACATTTACCTTGATAAGAATGAAACAGTGTACCAAGG
CATTCTCAATTCTGTGACACATACAAGGACAGTTAATATCTAATAATGAGTGGTGTAC
TGCTTCCAGAGACAGCAAGACCCACCTGCCAGACTGAGCTCAGCAATATTCAAGAGGCCA
AGGGTAAAGAACGCTCTAGGACAGTATATCCCCCTGTTGATGAAAGATGGTACTACAAGC
CAACACAATGTCATGGCAGTGTGGACAGTGTGGTGTGACAGATATGAAATGAAGTC
ATGGGATCCAGAATAAAATGGTGTGCAAGATTGTGCTATAGATTTGAGATCTCCGGAGATT
TGCTAGTGGGATTTCTAGAATGGACTGATGATGAGGAGATGATGAAGACGATATTGAAATG
ATGAGAATGAAATTGAAGATGATGATGAAGATGAAAGGGGTGATGATGATGGTGTGATGAC
CATGATGTATACATTTGATGATGACAGTGTGAAATCAATAATTCTACATTCTAATT
CAAAATGATAGCCTATTAAATTATCTCTCCCAATAACAAAATGATTCTAAACCTCA
CATATATTGTATAATTATTGAAAATTGCACTAAAGTTATGAACTTTATGTTAAAT
AAGAATCATTGCTTGAGTTTATATTCCCTACACAAAAAGAAAATACATATGCAGTCTA
GTCAAGACAAAATAAAAGTTGAAAGTGTACTATAATAAATTTCACGAGAACAAACTTTGT
AAATCTCCATAAGCAAAATGACAGCTAGTGCTGGATGTCAGTCAGTTAATTGGAAAG
ATAATTCTAAGTGAATTTAAATAAAATTGACCTGGCTTAAGGATTAG
AAAATATGATGCTTAAATGCAATTCAAAGTAGCATCTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCACATGACTCCAAAAAAAAAAAAAA

FIGURE 179

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829
><subunit 1 of 1, 436 aa, 1 stop
><MW: 49429, pI: 4.80, NX(S/T): 0
MLKVSAVLVCAAAWCSQSAAAAAVALAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFQALDPAKDPCLKMCKSRHKVIAQDSQTAVCISHRRLTHRMKEA
GVDHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCPCP
SDKPTSTSRRNVRKACSDLEPREVANRLRDWFKALHESGSQNKKTKTLLRPERSRFDTSLIP
CKDSLGLWMFNRLDTNYDLLLQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNNEWCYCFQ
RQQDPPCQTELSNIQKRQGVKKLLGQYIPLCDEDGYYKPTQCHGSVGQCWCVDRYGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDEDDEDDIMNDEDEIEDDDDEGDDDDGGDDHDVYI
```

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 180

FIGURE 181

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196
><subunit 1 of 1, 229 aa, 1 stop
><MW: 26017, pI: 4.73, NX(S/T): 0
MGDKIWLPPVLLAALPPVLLPGAAAGFTPSSLSDFTFTLPGQKECFYQPMPPLKASLEIEY
QVLDGAGLIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFSTISEKVIFFEL
ILDNMGEQAQEQQEDWKKYITGTDILDMLKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN
IQESNFDRVMFWSMVNLLVVVSAIQVYMLKSLFEDKRKSRT
```

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

09278514-101601

FIGURE 182

CCATCCCTGAGATCTTTATAAAAACCCAGTCTTGCTGACCAGACAAAGCATACCAAGAT
CTCACCAAGAGAGTCGCAGACACTATGCTGCCCTCCATGGCCCTGCCAGTGTGCTGGATG
CTGCTTCCCTGCCCTCATTCCCTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGA^{ACT}GCC
CTCTCCACGGATCAGCTGCCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCCTGT
TTTGTCA^{CCAAA}ATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTCCTCCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACATCTGGATTGGCTCCATGACCCCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGAGTGAGTAGCACTGATGTGATGAATTACTTTGCATGGAGAAAATCCC
TCCACCATCTAAACCC^TGGCCACTGTGGAGCCTGTCAAGAACAGGATTTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCC^TATGTCGCAAGTTCAAGGACTAGGCAGGT
GGGAAGTCAGCAGCCTCAGCTGGCGGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCC^TGGAAAGAGAATATTCTCCCCAAACTGCCCTACCTGACTACCTTGTCATGATCCTCC
TTCTTTCTCTTTCTCACCTTCAGGTTCTCTGTCTTCCATGTC^TGAGATC
TCAGAGAATAATAATAATGTTACTTTATAAAAAAAAAAAAAAA

FIGURE 183

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965
<subunit 1 of 1, 175 aa, 1 stop
<MW: 19330, pI: 7.25, NX(S/T): 1
MLPPMALPSVSWMLLSCLLQCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM
DADLACQKRPSGKLVSVLGAEGSFVSSLVRSISSNSYIWIGLHDPTQGSEPDGDGWEWSS
TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWDYNCDAKLPYVCKFKD
```

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

FIGURE 184

CCAGTCTGTCGCCACCTCACTTGGTGTCTGCTGCCCCGCCAGGCAAGCCTGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACCATGCGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGCTGGC
TGCGCTGCCGGAGAGCTGGCGCCGGCCCTGCGCTGCTACGGTCTGTCCGGAGCCCACAGGAGTGT
CGGACTGTGTCACCATGCCACCTGCACCAACGAAACCATGTGCAAGAACACTCTAC
TCCCGGGAGATAGTGTACCCCTTCAGGGGGACTCCACGGTGACCAAGTCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCCTGCCGTGTCTGCTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCGCTCTGAACAGCCTCCACTGGGGGCCCTCACGCTC
CTCCCACTCTTGAGCCTCCGACTGTAGAGTCCCCGCCACCCCCATGGCCCTATGCGGCCA
GCCCGAATGCCTTGAAGAAGTGCCTCTGCACCAGGAAAAAAAAAAAAAAA

FIGURE 185

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405
<subunit 1 of 1, 125 aa, 1 stop
<MW: 13115, pI: 5.90, NX(S/T): 1
MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP
FQGDSTVTKSCASKCKPSDVGIGQTLPVSCCNTELCNVDGAPALNSLHCGALTLLPLLSLRL
```

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGTCCCGGACCCCTGACTCTGCAGCGAACCGGC
ACGGTTTCTGGGGACCCAGGCTTGC~~AA~~GTGACGGTCATTTCTCTTCTCCCTCTT
GAGTCCTTCTGAGATGATGCTCTGGCGCAGCGGGAGCTACCCGGTCTTGTGCGATGG
TAGCGGCGGCCCTCGCGGCCACCCCTGCTGGGAGTGAGCGCCACCTGA~~ACT~~CGTTCTC
AATTCCAACGCTATCAAGAACCTGCCCCACCGCTGGCGCGCTGCGGGCACCCAGGCTC
TGCAGTCAGCGCCGCGCCGGAAATCCTGTACCCGGGGAAATAAGTACCA~~G~~ACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGACGCGAGGCGTGC~~AA~~CTGTCTCGCCTGCAGGAAGCGCGAAAAGC
CTGCATGCGTCACGCTATGTGCTGCCCGGAATTACTG~~AAA~~ATGGAATATGTG~~T~~TCTT
CTGAT~~AAA~~ATCATTTCGAGGAGAAATTGAGGAACCATCACTGAAAGCTTGGTAATGAT
CATAGCACCTGGATGGTATTCCAGAA~~G~~ACCACCTGTCTC~~AAA~~ATGTATCACACCAA
AGGACAAGAAGGTTCTGTTGTCTCGGT~~C~~ATCAGACTGTGCCTCAGGATTGTG~~T~~TGCTA
GACACTCTGGTCCAAGATCTG~~AA~~ACCTGTCTGAAAGAAGGTCAAGTGTG~~T~~ACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAA~~A~~ATTCCAGCGTTACTGTGGAGAAGGTCTGTC
TTGCCGGATCAGAAAGATCACC~~A~~CAAGCCAGTA~~TT~~CTCTAGGCTTCACACTTGT~~C~~AGA
GACACT~~AA~~AC~~C~~AGCTATCC~~AA~~ATG~~C~~AGTGAA~~CT~~CTTTATATAATAGATG~~T~~ATGAAAACC
TTTATGACCTTCATCAACTCAATCTAAGGATATA~~C~~AAGTTCTGTT~~C~~AGTTAA~~G~~CAT
TCCAATAACACCTTCCAAAACCTGGAGTGTAAAGAGCTTGT~~T~~TTATGGAAC~~CT~~CCCTG
TGATTGCAGTAAATTACTGTATTG~~AA~~TTCTCAGTGTGGC~~A~~CTTACCTGT~~AA~~TGCAATGA
AACTTTAATTATT~~TTT~~CTAAAGGTGCTGC~~A~~CTGCCT~~TTT~~CTCTGTTATG~~AA~~TTT
TTGTACACATTGATTGTTATCTT~~G~~ACTGACAA~~A~~ATTCTATATTG~~A~~CTGA~~G~~TAATCATT
TCAGCTTATGTTCTTAAAGCAT~~A~~CCCTTACCC~~C~~ATTAA~~T~~CTAGAGTCTAGAACGCA
AGGATCTTGGAA~~T~~GACAA~~A~~ATGATAGGTACCTAA~~A~~GTAA~~C~~ATGAAA~~A~~ACTAGCTTATT
TTCTGAA~~A~~ATG~~T~~ACTATCTTAA~~T~~ATTAT~~T~~CC~~T~~TTAGGCTGTGAT~~A~~CTTTG~~A~~
AAT~~AAA~~ATT~~T~~AA~~C~~ATT~~T~~AAAA~~AAA~~AAA

FIGURE 187

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530
<subunit 1 of 1, 266 aa, 1 stop
<MW: 28672, pI: 8.85, NX(S/T): 1
MMALGAAGATR V FVAM VAA ALGGHPLLGV SATLNSV LNSNAIKNL P PPLGGAAGHPGS AVSA
APGIL YPGGNKYQTIDNYQPYPCA EDEECGTDEY CASPTRG GDAGVQ ICLACRKRRKRCMRH
AMCCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVILKEGQVCTKHRRKGSHGLEIFQRCYC GEGLSCR IQ
KDHHQASNSRLHTCQRH
```

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

09078544-101601

FIGURE 188

TGTGTTCCCTGCAGTCAGAATTGGGACNGCAGGGTTCCCGGACCTGATTTGCAGCGGA
ACGGGAAGGTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTCTCTCCTTCNG
GAGTCCTNTGAGANGATGGTTGGGCCAGCGGGAGCTAACCCGGTTTTGTNGCGATG
GTAGCGCGGTTTCGGCGGCCACCTNTGCTGGGAGTGAGCGCCACCTGATCGGTTTC
AATTCCAACGNTATCAAGAACCTGCCCCACCGNTGGCGGCCTGCGGGCACCCAGGNTT
TGCAGTCAGCGCCGCGCCGGAAATCCTGTACCCGGGGAAATAAGTACCAAGACCATTGACA
ATTACCAAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGTGAGTA
CTGGCTAGTCCCACCCGCGAGGGGANGCGGGCGTGC
CAAATNTGNTNGCCTGCAGGAAGCGCCAAAACG
CTGCATGCGTCANGCTATGTGCTGCCCGGGAAATTACTG
CAAAATGGAATATGTG
CTGATCAAATCATTCGAGGAGAAATTGAGGAAC
CATCACTGAAAGCTTGGTAATGAT
CATAGCACCTTGGATGG

FIGURE 189

GAGGAACCTACCGGTACGGCCGCGCCTGGTAGTCGCCGTGCGTGCACCTCACCAATCCGTGCCGCCGG
CTGGCCCGTGGAGACTGGCTGCGCTCTCTCGCACGGGGTCTGGCTCGCAGGGGGTCCCGCCCCA
GGGGTTGGAGGGAGTGTACAGGAACGGGAGTGGCAAGGACCCCCGGATGGCAAGGATATTTTTGAGGAAAGGA
AGTATTAGAAATGAGCTGAGAACCTTACAGAATTAATTATTTGGGACAGATTGATGCTTGATGACCCCT
TGAAGTAAATGAGACAGAAAGTTCTCAAATTGCAATTACATCAACTGGAACCGACAGTGAATCTTAATGTCAC
TTAAATCAGAACCTTGCAATAAGAAAGAGAATGGGATCTGGTTAAATAGATGACTATATCAGAGACTTGGAAAAG
GATCATTCCTGTTCTGATAGTGTATAGGCCATTAGTGGCAGACAGTCAGGATTTCAGTTTACAGTTTACTGG
AGTGTCCAAAAGTGCAGAGTAAAGAACAGCTTCAGAAATGGGATTGAGTTTACATCTGTGATAAA
AAACCGGAATAACCAAATGCACTGGCAGATTTTAAAAATAAATAGAGCATATGAGTACTCAAAGATGAGA
TCTACGGAAAAGTATGACAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGCCAGTATGAAAGCTGGAA
CTTAAATCTGTTGATGTTGGTATTTAGTGTGATCTGAAATTCAAATCTGGAAAGAGAGAAATTGAGTGC
TCTGTTAATCTGGAGAACTGGTTGTTGAAATTTTACTTCCCAGCTGCTCACACTGCCATGATTAGCTCC
CACATGGAGAGACTTGGAGAAAGTGGGATCTGGTAAACTTGGAGACTTGGGATCTGGTAAACTTGGGATCTGG
GCTTGGCCGAATGAAAGGACTCAACGACTTCCAGCTCTTCTATTTCGGCTGGATGTTAGAAGTACAGTGACAGAACCTTG
TCATGGAGACAGATCAAGGAGAGTTTGTGAGTTTGCAGTCAGCATGTTAGAAGTACAGTGACAGAACCTTG
GACAGAAAATTGCTGAACTTCAACCAATTGGCTTGGCTGCTGTTGATCACCTTTGTTCTAA
AGGAGGAGATTGTTGACTTCAGACAGCTGGCTTGGCTGCTGTTGATGCTTCAACTTGTGAGTGTCAA
AGAAGATAATTGGAGAATTAACATAATCTTCCGAGATTGTTGAAACTTCTGGCAAAACACATAGAGGATGCTT
GGCTCATCTGGGGTGTGTTTCTATTGGGAAAAAAATGAAATTCAATGTCCTGAGCTGAAAGAAA
AAAAACTCTACTTAAATGATCATATTCAAGTTGGCAGGTTGACCTGTTCTCTGACCCAGACATCTGAGTAA
TCTGTTGATGTTTCTGGCTCTACAGCTTAAAGGAAAGGAAACAGGAAATGATGAAATTCTGATGAAA
GAAGATTCTATATGATATATCTGGCTTGGCCAAAAGGAAAGTGTGATTTCTGATGTTACACGCTTGACCTTCAA
TTTCTGGGATCTGGGAAAGGAAACCTTGGCTGTTGATGTTTCTGGCTGGCTGATCACCTTGGTCAA
ACCAAGATTGAGAACGACATCAATTCTTCTGTTGAGCTTGGCTGAGTACACTAGATTGATCAGTTGATG
GGGACTCTGTAACATGTTAATCAGGTTTCTCAAACACAGTGGTATTCAACAGCTTCAACATTCTGAGTAA
TGAAGGAGACTCTGTCGAAACATTCTGGAGTTCTAGGAGATTCTGAGGATTTGAACTCTTCTGAGTGGCTCCTCTAC
ACCCACCACTTCAACGAGACTGTTACACAAAGAAAACAACGAGACTGAGGTTGATTCTTCTGGCTG
GTGCTGATCTCCGGAAAGTATGGCAAAGTGGAAAAGATGGCCGGGAGCTTAACTGGGACTGATCACCTGGG
CAGTAATGATTGCCCACAGTATCATCTTCTTGTGCCCCAGGAACTGTTCAAGATACCTCTGAGCTT
TCCCGGAAATCAAATAAGCTTACAGTATGAGTCAATGTTGAGGATGGGAGCTTACCTTCTGAGGAAAT
CTGGGTCTGAGATTCTCTACAGTATCTCAAGGATCTTCAACAGCTTCAACCTCTGAGCTTCAACG
GAAAATCATGGGTGATTGATTCTATGCTCTTGGTGTGGACCTTGGCAAGAATTGCTCCAGAAATTGAGCT
ACCTGGCTAGGAGTAAAGGAAAGTGGAAAGGCTGGAAAGTGGCTGAGGCTTATGCTCAGACATGCCAGAA
AGCTGGGATCAGGGCTTCACTTCACTGTTGATTCTTCTGAGGAAAGGAGAAATTCTGAGGAGCA
GATAAACTACCAAGAGTCAACGCTGCTGCTTAAATGCTGAAATTGAGAAAATCTGGCAAAATCAGGCAA
GAGGAATAGGATGACTTGTATGTTGAGATGAGGAAAGGAAATCTGGCAAGGAACTTCTGAGGAGCT
AAGACACCTATTAGATGTTACATTATGATGGGAGTGAATGNAACATTATCTAGACTTGAGTTGACTGCCA
GAATTATCTACACGCTGGTAAAGGAGGGCTGCCAACTTCTGTTGAGGCTTAAAGGGGGTTTAAATTA
GACTTTGGCAGGCTTATAATAATGAGGTTTCTACACAGAACAGAATGAGTCTAGTATTCTTGTGTTGCT
TTTAAACACCTTAAATTAAGGAAACTGATTCTGAGGCTTACAGGAGCTTACAGGAGCTTACAGGAGCT
GACCATGATTGCTGCCCTCGACGGGACTTAAATGTTTCTGGGCTGCTGAGACATGAGTCTGCTGCT
ATCTACATATAATGTCAGTTGTTAAAGTCCACTTCCCTTCAGGTTTTGGCTGACCTGAAAAGAGGTA
TAGTTTTGGTCACTGTTCTTAAAGTCTTACAGGAAACATCTCCAACTTATTTCTGTTTAAACACCCAT
GATGTTGGCAGACTAAACAAACCTCTGTTATGCTGTTATTATGAGGAGATTCTCATGTTTCTTCTCTCA
AAGGTTGAAAAGGAAATGTTTAAATTCTGGGAGTGGCAAGCAGTGGCAGCACTTATGTCACACAGTA
AAATTGAGCAACAGTAAAGTCAGACAAATTCTGAGTTGCTGTTACATCAGGAAAACCTGAGGAAA
TAGCAATTAACTGGGATTGAGTATCTAAATGTTTACAGGTTTAAAGGTTTAAAGGAGTATA
TGTGTTGATGTTTCTGAAATGCTTCTGAGGAAATTTCCACAGGCTGAGGCTTACAGGAGCT
TTACATATACTGGCTCTGAGCTTGTGTTTACAGGTTTAACTGGGTTTCTTCTGAGTTTGG
AATGATACTGTGTTTATTCCAGTTACTAGTTACTGTCAGAGGGCTGCCCTTCTGAGAATAATTGACATA
ACTGAGGTTATTTTATAAGAAAATCAAGTATAAACTGAGGAAGGGATCTCTGAGTTCTGTTGTTT
CTCAAGGAGTCAAAATTGTCAGTAACTGAGTTGTTGTTGAGTAAATTCACTGAGTGTACAGAATGGTAAA
CCAATCAGTCAAAAGGGCTTGAATTAAAAGGGCTTGAACCTTTTCAAAAAAAAAAAAAAA

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439
<subunit 1 of 1, 747 aa, 1 stop
<MW: 86127, pI: 7.46, NX(S/T): 2
MGVWLNKDDYIRDLKRIIILCFLIVYMAILVGTQDFYSLLGVSKTASSREIRQAFKKLALKL
HPDKNPNNPNAHGDFLKINRAYEVLKDEDLRKKYDKYGEKGLEDNQQGQYESWNYYRYDFGI
YDDDPEIITLERREFDAAVNSGELWFVNPFSPGCSHCHDLAPTRDFAKEVDGLLRIGAVNC
GDDDRMLCRMKGKVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNS
IQATAFAAGIGWLITFCSKGGDCLTSQTRRLSGMLFLNSLDAKEIYLEVIIHNLPDFELLSAN
TLEDRLAHHRWLLFFHFHKGNENSNDPELKKLKTLLKNNDHIQVGRFDCCSSAPDICSNLVVFQP
SLAVFKGQGTKEYEIHGKIKLYDILAFAKESVNSHVTTLGPQNFPANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIHEYEGHHS
AEQILEFIEDLMNPSVVSLTPTTPNELVTQRKHNEWMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFFPPKSNSKAYQYHSYNGWNRDAYSRLRIWGLG
FLPQVSTDLPQTFSKVLQGKHNHWIDFYAPWCGPCQNFAPEFELLARMIKGKVKAGKVDC
QAYAQTCTQKAGIRAYPTVKFYFYERAKRNFQBEQINTRDAKAIAALISEKLETLRNQGKRNKDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

FIGURE 191

AGACAGTACCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAA
GCCATGAACATCATCCTAGAAATCCTCTGTTCTGATCACCATCATCTACTCCTACTTGG
GTCTGGTGAGATTTCTTCATTCCCTCAGAGGAGAAAATCTGGCTGGGAGATTGGTCTCA
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTGAATTGCAAAACGACAGAGC
ATATTGGTCTGTGGGATATAAAGCGCGGTGTGGAGGAAACTGCAGCTGAGTGGCGAAA
ACTAGGCCTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGGCCATCTCAGCACCAAGGATGAAGAGATTACCAAGACATTGAGGTCAA
CATCCTAGGACATTGGATCACAAAGCACTCTTCCATCGATGATGGAGAGAAATCATG
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCTTACCTCATCCATAT
TGTTCAGCAAATTGCCCTGTTGGCTTCACAGAGGTGACATCAGAACCTCAGGCCATT
GGGAAAAACTGGTATCAAACCTCATGTCCTGCCAGTTTGATGAATACTGGGTTACCCA
AAAATCCAAGCACAAAGATTATGCCCTGATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAATGATTTTGTTCATCGTATATCAATATCTTCTGAG
ACTACAGAAGTTCTCCTGAACGCGCTCAGCATTAAATCGTATGCAGAATATTCAAT
TTGAAGCAGTGGTGGCCACAAAATCAAATGAAATGAAATAAAGCTCCAGGCCAGAGATG
TATGCATGATAATGATATGAATAGTTCAATCAATGCTGCAAGCTTATTTCACATT
TCAGTCTGATAATATTAAAACATTGGTTGGACTAGCAGCAGTCAAACGAACAAGATTA
ATTACCTGCTTCTGTTCTCAAGAATATTACGTAGTTTCATAGGTCTGTTTCCCT
TCATGCCCTTAAACCTCTGCTTACATAAACATACTTAAAGGTTCTTTAAGATAT
TTTATTTCATTTAAAGGTGGACAAAGCTACCTCCCTAAAGTAAATACAAAGAGAACT
TATTTACACAGGGAGGTTAACGACTGTCAGTAGCATTCAAATCTGTAGCCATGCCACAG
AATATCAACAAAGAACACAGAACATGAGTCACAGCTAACAGAGATCAAGTTCAGCAGGCCAGCTT
ATCTCACCTGGACATATTAAAGATTCAGCATTGAAAGATTTCCCTAGCCTCTCCCTT
TCATTAGCCAAAACGGTGCAACTCTATTCTGGACTTATTACTGATTCTGCTTCTGTAT
AACTCTGAAAGTCCACCAAAAGTGACCCCTCTATATTCCCTCTTTTATAGCTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTAAATCTCAGAATTAAAGTTCTAGCCCCATGA
TAACCTTTCTTGTAAATTCTGCTTCATATATCCTGGTCCAGAGATGTTAGACAAT
TTTAGGCTCAAAATTAAGCTAACAGGAAAAGGAAGTGTACTGGCTATTACATAAGAAA
CAATGGACCCAAAGAGAAGAAA

FIGURE 192

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409
<subunit 1 of 1, 300 aa, 1 stop
<MW: 33655, pI: 9.31, NX(S/T): 1
MNIILEIILLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFRAKRQSI
LVLWDINKRGVEETAAECKLGVTAAHAYVVDCSNREEIYRSLNQVKKEVGDVTIVVNNAGTV
YPADLLSTKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYC
SSKFAAVGFHRLTSELQALGKTGIKTSCLCPVFVNTGFTKNPSTRLWPVLETDEVVRSLID
GILTNKKMIFVPSYINIFRLQLKFLPERASAILNRMQNIQFEAVVGHKIKMK
```

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 193

CGGGCGGCGCTCGGGCGCGAGGTGAGGGGCGGAGGTGAGGGGCGGAGGGTTCCACGCAGG
ATGCCCGGCTCTGCAGGAAGCTGAAGTGAAGAGGCCGGAGGGGCCAGCCCGCCCCGGC
AGGATGCCAAGGCCGGCTGTTCCGCTGTGGCTGGTGTGGGGTCGGTGTTCATGATCCT
GCTGATCATCGTGTACTGGGACAGCGCAGGCAGCGCAGCTTCACTTGACACGTCCTTCT
CTAGGCCACACGGGCCGCGCTGCCACGCCGGGCCGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTCGACGAGTTCTGGACAAGTTCTCAGTGTGGCTGAAGCAGAGCGA
CCTTCCCAGAAAGGAGACGGAGCAGCCGCTGCGCCGGGAGCATGGAGGAGAGCGTGAGAG
GCTACGACTGGTCCCCGCGCAGGCCGGCGCAGGCCAGACCAGGGCCGGCAGCAGCGGAG
CGGAGGAGCGTGTGCGGGCTTCTGCCAACCTCCAGCCTGGCTTCCCCACCAAGGAGCG
CGCATTGACGACATCCCCAACCTGGAGCTGAGCCACCTGATCGTGGACGACGGCACGGG
CCATCTACTGCTACGTGCCAACGGTGGCTGACCAACTGAAAGGCCGTGATGATCGTGTG
AGCGGAAGCTGTGCAACCGGGTGCCTACCGCAGCCATCCCGCCGAGCA
CGTGCACACGCCAGCGCACCTGACCTAACAAAGTTCTGGGCCGCTACGGGAAGCTCT
CCCGCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTCCTCTCGTGCACGGCACCCCC
TTCGTGCGCTGATCTCCGCTTCCGACGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTGCGCTGCCCCATGCTGCGCTGTACGCAACACACCAGCCTGCCGCTCGCGC
GCGAGGCCTTCCGCGCTGCCCTCAAGGTGCTTCGCCAACCTCATCCAGTACCTGCTGGAC
CCGCACACGGAGAACGCTGGGCCCTCAACGAGCACTGGCGCAGGTGTACCGCCTTGCCA
CCCGTGCCAGATCGACTACGACTTCGTGGGAAGCTGGAGACTCTGGACGAGGAGCCGCGC
AGCTGCTGCACTACTCCAGGTGGACCGGCAGCTCCGCTCCCCCGAGCTACCGAACAGG
ACCGCCAGCAGCTGGAGGAGGACTGGTTCGCCAACATCCCCCTGGCTGGAGGAGCAGCT
GTATAAAACTCTACGAGGCCGACTTTGTTCTTCCGGTACCCCAAGCCGAAAACCTCCCTCC
GAGACTGAAGCTTCGCGTTGCTTTCTCGCGTGCCTGGAACCTGACGCACGCCACTCC
AGTTTTTTATGACCTACGATTTGCAATCTGGCTTCTGTTCACTCCACTGCCCTATCC
ATTGAGTACTGTATCGATATTGTTTAAGATTAATATTTAGGTATTTAATACGA

FIGURE 194

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112
<subunit 1 of 1, 414 aa, 1 stop
<MW: 48414, pI: 9.54, NX(S/T): 4
MTKARLFRLLWLVGSVFMILLIIIVYWD$AGAAHFYLHTSFSRPHTGPP$PPTPGPDRDRELTA
DSDVDEF$LDKFLSAGVKQSDLPRKETEQPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSSLAFPTKERAFDDIPNSELSHLIVDDRHGAI$CYVPKVACTNWKRVMIVLS
GSLLHRGAPYRDPLRIPREHVHNASAHLTFNKFWRRYGKLSRHLMKV$KLKKYTKFLFVRDPF
VRLISAFRSKFELENEEFYRKFAVPM$LRILYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPFNEHW$RQVYRLCHPCQIDYDFVGKLETLD$EAAQ$QLLQ$LLQVDRQLRFPPSYRNRT
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLFGYPK$PENLLRD
```

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

FIGURE 195

TCGGGCCAGAATT CGGCACGAGGCGGCACGAGGGCGACGGCCTCACGGGCTTGAGGTGA
AAGAGGCCAGAGTAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGGAACCGCCT
ATGCCGGAAAGGTGGTGGTCGTGACCGGGGCGGGCGGCATGGAGCTGGATCGTGC
GCCTTCGTGAAACAGCGGGGGCCGAGTGGTTATCTGCAAGGATGAGTCTGGGGCCGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCTCTGGTTCTGAGACCATCCGCCGATTGGCCGCTGGATTGTGTTGTCAACAAAC
GCTGGCCACCACCCACCCCCACAGAGGCCGAGGAGACCTCTGCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACTACTGGGACGTACACCTTGACCAAGCTGCCCTCCCTACCTGCCGA
AGAGTCAGGAATGTCATCAACATCTCCAGCCTGGTGGGGCAATCGGCCAGGCCAGCA
GTTCCCTATGTGGCCACCAAGGGGCAGTAACAGCCATGACCAAAAGCTTGGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCCAGGAAACATCTGGACCCGCTGTGGG
AGGAGCTGGCAGCCTTAATGCCAGACCCCTAGGGCCACAATCCGAGAGGGCATGCTGCCAG
CCACTGGGCCGATGGGCCAGCCGCTGAGGTGGCTGCCAGTGTCTGGCCTCCGA
AGCCAACCTCTGCACGGCATTGAACTGCTCGTGA CGGGGGGTGCAGAGCTGGGTACGGGT
GCAAGGCCAGTCGGAGCACCCCGTGGACGCCCGATATCCCTCCTGATTTCTCATTT
CTACTTGGGGCCCTTCCTAGGACTCTCCACCCCAAACTCAAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAACGCCAGTTAGCAAGGTGCCGGTCACCCCTGCAGGTTCCCAT
AAAAACGATTGCAGCC

FIGURE 196

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045
<subunit 1 of 1, 270 aa, 1 stop
<MW: 28317, pI: 6.00, NX(S/T): 1
MATGTRYAGKVVVVTGGGRIGAGIVRAFVNSGARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRFGLDCVNNAGHHPPPQRPEETSQGFRQLLELNLLGTYTLLTKL
ALPYLRKSQGNVINISSLVGAIGQAQAVPYVATKGAVTAMTKALALDES PYGVRVNCISPGN
IWTPLEELAALMPDPRATIREGMLAQPLGRMGQQPAEVGAAAVFLASEANFCTGIELLVTGG
AELGYGCKASRSTPVDAPDIPS
```

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

FIGURE 197

AGGGGGGCGAGCTGCAGGCTGACCTTCAGCTGGCGGAATGGACTGGCCTCACAAACCTG
CTGTTCTTACCATTCCATCTTCTGGGCTGGGCCAGCCCAGGAGCCCCAAAGCAA
GAGGAAGGGCAAGGGCGCTGGGCCCTGGCCCTGGCCCTCACAGGTGCCACTGGACC
TGTTGTCACGGATGAAACCGTATGCCCATGGAGGAGTATGAGAGAACATCGAGGAGATG
GTGGCCAGCTGAGGAACAGCTCAGAGCTGGCCAGAGAAAGTGTGAGGTCAACTTGAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGCTACAGCATCAACCACGACCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCTGTCTGGGCTGTGAACCCCTTCACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGTGTTCAGCCAGGTTCTGTGCGCCGCCG
CCTCTGCCGCCACCGCCCCCAAGGCCCTTGCCGCCAGCGCGAGTCATGGAGACCATCG
CTGTGGCTGCACCTGCATCTTCTTGAATCACCTGGCCAGAAGCCAGGCCAGCAGCCGAGA
CCATCCTCCTTGACCTTGTGCCAAGAAAGGCCATGAAAAGTAAACACTGACTTTGAAA
GCAAG

FIGURE 198

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294
<subunit 1 of 1, 180 aa, 1 stop
<MW: 20437, pI: 9.58, NX(S/T): 1
MDWPHNLLFLLTISIFLGLGQPRSPKSKRKQGRRPGLAPGPHQVPLDLVSRMKPYARMEEY
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSIHDPRIPVDLPEARCLCL
GCVNPFMEOEDRSMVSVPFSOPVVRRLCPPPPRTGPCRORAVMETIAVGCTIF
```

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

FIGURE 199

GCGCCGCCAGCGTAGGGGGTGGCCCTTGCCTCCGCTTCTGAAAAACCGGGCGGG
CGAGCAGGCTCGGGGCCCGCTGCCCTCCCCACTCCCGCCAGAAGCCTGCTG
GCGCCCAACATGGCGGGTGGCGCTGCCCGCAGCTAACGGCGCTCTGGCCGCTGGAT
CGCGCTGTGGCGCAGGCCAGGCCAGGAGGCCGCTGCCGCCAGAGAGGCCGG
TCCAGCCCCATGGTGTCCATCTGCCAGACTGATTCAAGAATGGGAGCTTTGAAAGAA
TGGTAAATACTCAGATCAGTGTGGGAAGGTAGATGTCATTCAAGAACCAGGTTGAGTG
GCCGCTTCTTGTCAACACTCTCCAGCATTTTCATGCAAAGGATGGGATATTCCGCCGT
TATCGTGGCCCAGGAATCTCGAAGACCTGCAAAGATTATCTTAGAGAAGAAATGGCAATC
AGTCGAGCTCTGACTGGCTGAAATCCCAGCTCTAACGATGTCAGGATGGTC
TTTTAGCATCTGGCAAGATATGGCATCTCACAACTATTCACAGTGA
CTCTGGCTTGGTTCTTATGTGTTCTGCATAGCCACCTTGGTTTGGCTTTTATGGG
TCTGGTCTTGGTGTAAATCAGAATTTCTATGTGCCACTTCCAAGGCATTATCTGAGC
GTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAGCTGAAACAGTTGCA
GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
GAAAAGAAGATCTGGCGATGAGGATGAAAGAAAACAAAGCAGGCCCTTGTAGATG
GAAAGAAGATCTGGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
CTGGTGTGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
ACCCGGAGGAAGTAGAGGCTGAGGAGGCTGAGAAGGAGGAGGAGGAGGAG
TGACACAGAGGTGGAAAGACTCTTGTAGGGCAGCGTAAAGTCAGCATG
TGAGATTAAATGATGCTTCAAGAATACACACCAAAACATATGTCAG
CCTGCAGTTGTACCAAATCTTAATTTCCTGAAATGAGCAAGCTCTTAAAGATGCT
CTCTAGTCATTGGTCTCATGGCAGTAAGCCTCATGTTACATAGGAGAGTCTTCAGGTG
GACAATCAGGATATGAAAACAAACGAGTGTGGGATCTGTTGGAGACTGGGATGGGAA
CAAGTCATTACTTAGGGTCTAGAGACTCTGAGGAGGCTGAAATGAAAGCC
AGCACCTTCAGAGACAAGGCTGAGGCCCTGTGAAATGAAAGCC
CTTGAGCATCCCCAAAGTCAACGTAGAGGCTTGCATCTTCTGTG
TTTGTCAAAATGCAAGGAAACATCAGGCACACAGTGCATG
ATTGAAAGGGCTTGGGTATAGAGAGCAGCTCAGAAGTCATCCAG
TGCTATGTTTATCTAACCTTAAATTTCAGCATTTCCACCATGGCATT
CCACACTCTTCACTATTATCTTGTGAGGAGACTC
TGTGTTGTTCTGACCTAACGGGTTTAGATAATCAGTAAC
GACTGCCAAACATCTCAATGAAATGTGTGGCATCAGAGACT
TTACAAGACAGATTAAAAAAATTTGTCAAAATATAGTTGTTG
AAGTTTCTAAGCAATTTCAGGCCAGACTCTAACG
CTTGTGAAGAAAAGTGAATACTGTTTGTGTTTCA
TACTTAAATAACTAAAACCACCTCTGATTTC
GAATTAAATGAACTCCAGTAACCTGAAAGT
ITCCTAAAGAATTATCTTGTAAATCTCA
GCTAATTCTT

FIGURE 200

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433
<subunit 1 of 1, 349 aa, 1 stop
<MW: 38952, pI: 4.34, NX(S/T): 1
MAGGRGCPQLTALLAAWIAAVAAATAGPEEAALPPEQSRVQPMTASNWTLMEGEWMKFYAP
WCPSCQQTDSEWEAFAKNGEILQISVGKVDVIQEPLGSGRFFVTTLPAFFHAKDGIFRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWNKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFFVIATLVFGLFMGLVLVVISECFYVPLPRHLSERSEQNRRSEEAHRAEQLQDAEEEK
DDSNNEENKDSLVDDEEEKGLEDLGEDEAEAAAAEDNLAAVGDEERSEANDQGPPGEDGVTR
EVEPEEAEGISEQPCPADTEVVVEDSLQRKSQHADKGL
```

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

FIGURE 201

ATCTGGTTGAACTACTTAAGCTTAAATTGTTAAACTCCGGTAAGTACCTAGCCCACATGATT
TGACTCAGAGATTCCTTTGTCACAGACAGTCATCTCAGGGCAGAAGAAAAGAGCTCC
CAAATGCTATATCTATTCAAGGGGCTCTCAAGAACAATGGAATATCATCCTGATTAGAAAAT
TTGGATGAAGATGGATAACTCTAACATTCACTTCGACTCTCAAAGCAATACCAGGATAGCTGT
TGTTTCAGAGAAAGGATCGTGCATCTCCTCTGGGTACCATGGGGTTCTTCC
AGCCCTGTCCTCTAATTGGATTATATGAGAAGAGCTGTTATCTATTCAGCATGTCACT
AAATTCTGGATGGAAGTAAAGACAATGCTGGCAACTGGGCTCTAATCTCTAAAGATAG
ACAGCTCAAATGAAATTGGGATTATAGTAAACAAAGTGTCTCCCAACCTGATAATTCTATT
TGGATAGGCCCTTCGGCCCCAGACTGAGGTACATGGCTCTGGAGGATGGATCACATT
CTCTCTAATTTCAAGATCAGAACACAGCTACCCAGAAAACCCATCTCAAATTGTG
TATGGATTCACTGTCAGTCATTGAGCAACTGTGTAGTGTGCCCCATATAGTATTGT
GAGAAGAAGTTTCATGTGAGGAGGGTGGAGAAGGAGAGAAATATGTGAGGTTAGTA
AGGAGGACAGAAAACAGAACAGAAAAGAGTAACAGCTGAGGTCAAGATAATGCAGAAAATG
TTAGAGAGCTTGGCCAACGTAATCTAACAGAACAGAAAAGGGAGGGCTGTGATTCT
GTATTGTGCACTCAAGGTAGGCTAGTATTATTTCTAGTTAGTAGATCCCAGACATGG
AAATCAGGGCAGCCAAGCTTGAGTTTATTATTTATTATTTTTTGAGATAGGGTCT
CACTTGTGTAACCCAGGCTGGAGTGCAGTGGCACAATCTGACTCACTGAGCTATCTCGC
CTCAGGCCCTCAAGTAGCTGGGACTACAGGTGCACTGGCACCAGCTTAATTGGTG
TTTTTGAGAGACTGGTTTGCATGGTACAGGCTGTTGACCAAGCTGTTCTACTCCTGGCTTAAG
TGATCTGCCGCCCTGGCTTCCAAAGTGTGGGATTACAGATGTGAGGCCACACCTGGC
CCCAAGCTGGAATTTCATCTGCATTGACTTGCATTTACCTTGCTTAAGCCATAAGCGA
ATCTTAATTCTGGCTCATAGTGTGCTGTCACAACATGCCATTGAAAGTCACGGT
GTGTTGCCACGATTGACCTCAACTTCAGCACTGTTAGTATCAGTTAGTAACTGAGGGTGA
ATATTCTGAATAGCTAAATGAGAACATGGAAAAAAATCTCACCAACAGTCAGAGCAATT
ATTATTTCTCATGATGATCATTAATTGATTATCATCTTAGTAAAGCAGGAACTCTA
CTTTTCTTATCAATTAAATAGCTCAGAGACTACATGCCATATCTCTAATGAAATCTT
TTTTTTTTTTTTTTGAGACAGAGTTCGCTTGTGCCCCAGGCTGGAGTGCACCG
CACGATCTGCCCTACCCGAACTCCCGCCCCCTGGGTAAGCAATTCTCGCTCAGCCT
CCCAAGTAGCTGGGATTACAGTCAGGCCACACCAACCCGGCTAATTGGTATTGGTAGT
AGAGACAGGGTTCTCCATGTCGGTCAGGGTAGTCCGAACCTCTGACCTCAAGTGTACTGC
CTGCCCTGCCCTCCCAAGTGTGGGATTACAGGCGTAGGCCACTGCAACCCAGCTAGAATCT
TGATAAATGTAATTGTAAGGAAACTGCTCTCATAGGAAGTTCTGCTTTAAATACA
AAATACATAAAATCATAAAATCTGATGATGATATAAAAAGTAACCAACCTCATTGGA
ACAAGTATTAAACATTGGAAATGTTATTAGTTGTGATGTAAGTGTCTTACATT
ACCATTGGTCAATTACTGTAAGGTTATTAGGAAACTGAAACTATATTCTCATG
TGCTGATTGTCTTATTCTCATCTTCCACTGGCTCATTTTATTCCATG
TTCTGTATTACTAGGGAGGCATTACAGTCCTCTAATGTTGATTAATGTA
TGTACCAATTACTAAATTATGAGTTAAATGGATGTTATGTTATGTTGAGTT
TTCATAAAAAAAAAACTTATCAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 202

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912
<subunit 1 of 1, 201 aa, 1 stop
<MW: 22563, pI: 4.87, NX(S/T): 1
MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSCAASPPWRLIAVILGILCLVILVIAV
VLGTMGVLSGPCPNWIYEKSCYLFSMSLNWDGSKRQCWQLGSNLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM
```

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

FIGURE 203

FIGURE 204

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921
<subunit 1 of 1, 693 aa, 1 stop
<MW: 77738, pI: 8.87, NX(S/T): 7
MTPQSLLQTTLFLLSLLFLVQGAHGRGHREDFRFCQRNQTHRSSLHYKPTPDLRISENSE
EALTVHAPFPAAHPASRSFPDPRGLYHFPLYWNRHAGRLHLLYGRDFLLSDKASSLLCFQH
QEESLAQGPPLLATSVTSWSPQNISLPSAASFTFSFHSPPHTAAHNASVDMCEKLKDLQLL
SQFLKHPQKASRRPSAAPASQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTLFQRTKGRSGEAEKRLLLVDFFSSQALFQDKNNSQVLGE
KVLGIVVQNTKVANLTEPVVLTQHQQLQPKNVTLQCVFWVEDPTLSSPGHWSSAGCETVRRE
TQTSCFCNHLTYFAVLMVSSVEVDAVHKHYLSLLSYVCVVSALACLVTAAYLCSRVPLPC
RRKPRDYTIKVHMNLLAVFLLDTSFLLESEPVALTGSEAGCRASAIFLHFSLLTCLSWMGLE
GYNLYRLVVVEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVDNYGPIILAVHRTPEGVIY
PSMCWIRDSLVSYITNLGLFLSVFLFNAMMLATMVQILRLRPHTQKWSHVLTLLGLSLVLG
LPWALIFFSFASGTFQLVVLYLFSIITSFQGFLIFIWYWSMRLQARGGSPPLKSNSDSARLP
ISSGSTSSSR1

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

FIGURE 205

TGCCTGGCCTGCCTTGTCAACAATGCCGTTACTCTGCTTCCAGGTTGCCCTGCCCTGCAGA
GGAAANCNTCGGGACTACACCNTCAAGTGCACATGAACCTGCTGCTGGCCGTCTTCCTGCTG
GACACGAGCTTCCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA
GCCAGTGCCATCTTCCTGCACTTCTCCTGCTCACCTGCCCTTCCTGGATGGGCCTCGAGGGG
TACAACCTCTACCGACTCGTGGTGGAGGTCTTGGCACCTATGTCCCTGGCTACCTACTCAA
GCTGAGCGCCATGGGCTGGGCTTCCCCATCTTCTGGTGACGCTGGTGGCCCTGGTGGATG
TGGACAACTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCC
TCCATGTGCTGGATCGGGACTCCCTGGTCAGCTACATCACCAACCTGGCCTTCAAGCCT
GGTGTTCCTGTTCAACATGG

FIGURE 206

CGGACGGCTGGCGGACCGCTGGGCAGCGCTGGGCGAGCGTGGCTGGTCAGGTCCAGGTTTGCTTGA
TCCTTTCAAAACTGGAGACACAAGAGGGCTTAGGAAAATTTGGATGGATTATGGAAACTACCCCT
GGCATTCCTCTGGCCAGGCCTCGGCTTCCACCCAGTGAGCTCAGGCTTCCCTGGCGGTGGTAAGAGAC
TCGGGACTCGCTGCTTCAAAGTCCCCCGTGAGTGAGCTCAGGCTTCCCTGGCGGTGGTAAGAGAC
TTCTCCCTGCTGACATCTGCCCTGGCGGAGAGACAGGGACTCAGGGGAATCCAACCTGAGTAGTAAATTCC
AGTTTCCAGCJAACAGGAACAGCAGGAGTACAGAGTCTCAGGAGAATTTACTGTCTACTAATG
GAAGTATTACAGCCAAAGGTTCTCATACTTACAGGAATACCGGCTTGGATGGAGATTACTAGCAGTAG
AGGAAATGTATGGATAACACTTACAGTTGATGAAGGACCCAGTGAGACTATATTAGGGCTCTGGTACTGTACCAG
ATGATTTTGAGGAAGTTGAGGAACCCAGTGAGACTATATTAGGGCTCTGGTACTGTACCAG
GAAAACAGATTCTAAAGGAATCAAATTAGGATAAGGATTGATGAATATTCTCTGAACAGGGT
TCTGCATCCACATCACACATTGCAACTTACAGGAGTGGCTTCAAGGACTTACGTTGAGACCTTATTGATATCTGAACAG
TGCCACTGGACCTGCTTAAATGCTATAACTGGCTTACTGCTTACGTTGAGACCTTACGTTGAGAC
AGAGATGGCAGTTGGACTTAAAGAGATCTAGGCAACTTCTGGCAAGGGCTTGGTTTTGGAA
GAAAATCAGCTGGTGGATCTGACCTCTAACAGAGGGTAGAGATTATACAGCTGACACCTCTGAACTCT
CACTGTCATAAGGGAAGAACATTAGAGAACCCGATACCATTTCTGGCAGGGTGTCTCTGGTTAAACGCTGTG
GTGGGAACCTGTGCTGTGCTTCCACAAATTGCAATGATGTCAAGTGTCTGGCAAGGAAAGTTTACTAAAAAAATACC
ACGAGGCTCTCAGTTGAGACCAAAGACGGCTGTCAAGGGGATTGCAAAACACTCACCCGACGTGGCCCTGGAGC
ACCATGAGGAGTGTGACTGTGAGAGGACACAGGAGGATGGCGCATCACACCCAGCAGCTTGTGCCA
GAGCTGTGAGGTGTGCTGTGCTTATTAGAGAACCTGTGCTTATCTGTTTACTCTCATCTTAAATTCTGACTGTGCT
TCAGGACCTTCTCATCTCAGGATTACAGTGTGCTGAAAGAGAGACATCAACAGAAATTAGGAGTGTGCA
ACAGCTCTTGTGAGAGGAGCTAAAGGACAGGAAAAGGTTCTCAATCTGGAAAGAAAATTAAATGTTGAT
TAATAGATCACCGCTAGTTGAGCTTACAGTTTCACTAGCTGGTTCTGTTTCTGTTTCTGCTTCTTC
GATACTGGCTTAGGGTATGTGACTACAGGAAAGAAAACCTGTCAAGGTGACCTCTGCTTGTGCTTGTGCTTAC
TCTAAAGCTCATGTCTGGCTTAAATCTGTTTAAATTCTGTTTCTGTTTCTGCTTATTCATCACAT
ATGTAACACAGAACATTCTATGTAACACAAACCTGGTTTAAAGGAAACTATGTGCTATGAAATTAAACTGT
GTCACTGATGAGGACAGACTGGATTTCTATTTCTGTTTCTGTTTCTGTTTCTGTTTCTGCTTATTCATCACAT
TTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT
TTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT
AAACAAATTGTTATGGCAGAGGAAACAAAGATGATATAAAATATGTTGCTCTGACAAAAATACATGTT
TCTGCTTATGGTCTAGAGTTAGGATTAATCTGCTTAAAGGAAACTTACGTTGAGGATTAAGATTGGTAAGTTGCAA
GACTTTTGAAATAATTAAATTATCATATCTTCCATTCTGTTTCTGCTTATGGAGATGAAAATAAAACCACTTATGA
AAGTAGACATTCTAGATCCAGCTCAGGCTTACTAACCTTCTTCTGGGAAATCTGACCTTAGCTCAGAAAACAT
AAAACACCTTGTGAAAGAACATTGGCAGCTTCTCTGATAAAGGGCTGTGCTGTGCTGAGTACACATCTATTAA
TTGTGATGTGTGTTTATTTATCTTAAACTCTGTTCCATACACTGTGATATAAAATACATGGATTTTTATGTACA
GAAGTATGTCTCTAACAGTTACTTATGACTCTGGCAATTAAAGGAAATCAGTAAAATATTGCTGT
AAAATGCTTAAATATNGTGCCTAGGTTATGTGTTGACTATTGAAATCATCAAAATGTTGATTAAGGATAACAGGGTAATGGCC
ATGTGGCTTATTGTTGGGAGAAAATTAAAAAAAGGGTTAGGGATAACAGGGTAATGGCC

FIGURE 207

MSLFGLLLTSLAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLWWRVLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPGKQISKGNQIRIFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFFVFRKSRVVDLNLLTEEVRLY
SCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVP SKVTKKYHEVLQ
LRPKTGVRGLHKSLTDVALEHHEECDCVCRGSTGG

Signal sequence:

amino acids 1-14

FIGURE 208

CCCATCTCAAGCTGATCTGGCACCTCTCATGCTCTCTTCAACAGACCTACATTCCATTGGAAAGA
AGACTAAAAATGGTTTCTCAATGTGGACACTGAAGACAAATTCTTATCCTTTAACATAATCTATTTC
AAACTCTTGGGCTAGATCTGGCTTCTAAACTCTGGCCTGTGAGTCACTCTGGATGTTCAAAGAACCCATGTG
ATCGTGACTGCAACAGAACGACTTCCAGAACGAACTTCTGGAGGTATTCCACAGAACACCAACGAACTCACCCTC
ACCATTAACCAACATACAGACATCTCCCAGCGTCTTACAGACTGGACCATCTGGTAGAGATCGATTTCAGA
TGCACACTGTGACTTATCTGGGCTGAAACAAACAGTGTGATCAAGAGGCTGAGAATTAAACAGAACGAG
TTAGTGACTCTATTAAACTCCCTTACCTGGATGAAACAGACTAGAGATAACCGAGGCTCCCG
CCCTAGCTTACAGGCTTCAGCCTGGGCCCCAAACTGTATTATCGAAACATCTGTGATTTCTATTAATAGAGAAA
GATGCCCTTCTAAACTTGACAAGGTTAAAGTGGCTCTCCCTGAAAGATAACATGTACAGCGTCCCTACTGTT
TTGCGCATCTACTTTAACAGAACTTATCTACAAACACATGTGAAAGATAACAGAACGAGGATGATTAAAC
CTCAACCAATTACAATTCTTGACCTAAGTGAAATTGCCCCCTGTGTTATAATGCCCCATTCTTGCGCC
TGTAACCTTCTGAAACTTCTGGGCTACAGATCCCTGATGCGCTGACAGAAATTAAAGGTTTACGCTTA
CACAGACTCTCTCAGCATGTGCCCCCAAGATGTTAACACATCAAAACACTCCAGGAACCTGGATCTG
CAAACATTCTGGCCAAAGAAATGGGGAGTGTGAAACATCTGGCATCTTCTGAAAGGAGCTG
TCTTCAATTGTTAACCTCTGGGATCTATGTCATGATCTACAGGAACTTCTGGGCTTAACTATCTGG
AAATTCTCGCGATCAAGAGGATATGCTTAAAGGATGTGAAAGGCTTAACTCTGGGCTTAAACATTAAAGA
ATCTTGAAGTCTTCTGACTTCTGGCAACTTCTTAAATGGTCAACCTGGCATGTTAAACAAATTAAAGA
CTGAGATCATGATCTTCTGAGTAAATAATCACCTTCAAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAT
GCCAGAACTCTGTGAAAGGTTGACCCCCAGCTTGGAAACAACTTCAATTATTCAGATATGATAAGTATGCA
AGGAGTGTGAGATTCAAACAAAGGGCTCTTCTGTTATGCTGTTAATGAAAGGCTGACAGAATTGGCAGACC
TTGGATCTAAGTAAAGTAGTATTTTGTCAAGTCTCTGATTCTCAGCATCTTCTCTCAAAATGCTG
AACTCTGAGGAAATCTCATGGCAACTCTTAAATGCGACTTCTGAAAGGCTTACAGAGGCTGAGATATTG
GACTTCTTCAAACCCGGCTGATTCTTCACTTCAACAGCATTTGAAGAGCTTCAAACTGGAAAGTCTGG
ATAAGCTGAGTAAATGGGATTATTCTCAATCGAGGAAATTCTCAATGCTAACCTTACAGGAACTTAAAGGTT
CTCGAGAAACTGATGTGACGACAATGACATCTCTTCTCCACAGCAGGACATGGAGACTGAGTCTCTTAA
ACTCTGGAACTCAGAGGAATCAGTAGTGTGTTATGGAGAGAAGGTGATAACAGATACTAACATTAACT
AACTCTGCTAAATGGAGGAAATTAGACATCTCTTCAAAATTCCTCAAGTGTGTTCTGGGAGTTTGTG
ATGCGCTTCAAATCTTAAAGGATCTCTTCTGGCCAATGGCTCAAACTTCTGAGTGGAGAAACTCCAGTGT
CTAAAGAACCTGGAAACTGGACCTGGCCACCAAACTGGGACTCTGCTCAGAGGATTATCTAACGTT
AGAAGGCTTCAAGAATCTGATTCTTAAAGAATAACCTGGGACTCTGCTCAGAGGATTCTAACGATGCTTC
CAGTTGGGATATCTGGATCTCAGCTTAAATGGAGATGATCCTAAAGGACCGAGCTTCTCCAGAAATGTC
AAACATCTGAGGATGTGCTTTCATGCTATAATGGTTCTGCACTGAACTCTGAGTGTGTTCTGGG
GTAAACCATACGGAGGTGACTATCTCTTACCTGGCACAGATGAGTGTGAGTGGGGCCAGGAGCACAAAGGG
CAAAGTGTGATCTCCCTGGGACTCTGACACCTGTGAGTTAGATCTGACTAACTGATTCTGTTCTACCTTCC
CTGTTATCTCTCTTCTCATGGTGTGATGACACAGACTCCTCTTCTGGGATGTGGTATATTCTAC
TTCTGTTAAAGGCAAGATAAAGGGGATCTCGCGTCTAATATCAGGACTGTTGAGTGTGTTTATCTG
GACACTAAAGGACCCAGCTGTGACGGAGTGGCTGAGCTGCTGGGCAAAACTGGAAAGACCCAAGAGAGAAA
CAGCTTAAAGGAAAGAGCAGTGTGTTGAGACAGAACGACTGCAAAGACTGAAAATTAAAGATAGCATT
TTGTCCTCATCGAGGCTCATGGTGTGAAAGGTTGATGTTCTGATTAATCTGATATTCTTCTGAGAAG
TCCAAAGTTCTCCAGCTCGGAAAGGGCTCTGTTGGAGTTCTGCTCTGAGTGGCCACAAACCCGCAAGCTCAC
CCATACTCTGGCAGTGTCTAAAGAACGCCCTGGCCACAGACAACTGAGGCTTATAGTCAGGTGTTCAAGGAA
ACGGCTAGCCCTTCTTGCACAACTGCCTAGTTACCAAGGAGAGGCCCTGGC

FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCTDVTLDVPKNHVIVDCTDKHLTEIPGG
IPTNTTNLTINHIDPDISPASFHRLDHLVIEIDFRNCVPIPLGSKNMCIKRLQIKPRSF
GLTYLKSLSYLDGNQLLIEPQGLPPSLQLSLEANNIFSIRKENLTELANIEILYLQCNQCYR
NPCYVSYISIEKDAFLNLTKLKVLSSLKDNNVTAVPTVLPSLTLYLYNNMIAKIQEEDFNNL
NQLQILDLSGNCPRCYNAPFFPCAPCKNNSPQIPVNADFTELKVLRLHSNSLQHVPPRWF
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELOVYRASMNLSQAFSSLKSL
KILRIRGYVFKELKSFNLSPLHNLQNLLEVLDLGTNFIFKIANLSMFQFKRLKVIDL SVN KIS
PSGDSSEVGFCNSARTSVESYEPQVLEQLHYFRYDKYARS CRFKNKEAS FMS VNE SCY KYGQ
TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAE LRYLDFSNRNL DLLH
STA FEELH KLE VLDI SSNS H YF QSEG ITHMLNFTKVLQKL MMND NDI SSTS RTMESES
LRTL EFRGNH LDVL WREGDN RYLQLFKNLLKLEEL DISKNSL SF LPSGVFDGMPPNLK NLSL
AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVPERLSNCRS LKNL LKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQM IQKTSFPEVN LNLKMLLHHNRFLCTCD AVWFVWWVNHTEV TIP
Y LAT DVT C VGP G A H KG Q S V I S L D LYTC ELD L TN L L F S L S I S V S L F L M V M M T A S H L Y F W D V W
Y I Y H F C K A K I K G Y Q R L I S P D C C Y D A F I V Y D T K D P A V T E W V L A E L V A K L E D P R E K H F N L C L E E
RDWLPGQPVLNLSQSITLSKKT V F V M T D K Y A K T E N F K I A F Y L S H Q R L M D E K V D V I I L I F L E
KPFQKSFKFLQLRKRLCGSSVLEWPTNPQAH PY FWQ CLK NAL ATDNH V A Y S Q V F K E T V

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 210

FIGURE 211

MENMFHQSSMLTCIFLLISGSCELCAEENFSRSYPCDEKKQNDSVIAECSNRRLQEVPQTGV
KYVTELDLSDFNIFITHTNESFQGLQNLTKINLNHNPNVQHQNGNPGIQSNGLNITDGAFLNL
KNLRELLLEDNQLPQIPSGLPESLTELSLIQNNIYNITKEGISRLINKNLYLAWNCFNKV
CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLINL
TLLDLSGNCPRCFNAPFPCVPCDGGSASINIDRFAFQNLTLRQLYLNLSSTSCLRKINAAWFKNM
PHLKVLVDLEFNLYLVGETIVSGAFLTMLPRLIEILDLSFNYIKGSYPQHINISRNFSKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINFQKQIDFKLFQNFNSNLEIIYLSENRISPLV
KDTRQSYANSSSFQRHIRKRRSTDFFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACLNLSANSNAQVLSGTTEFSAIIPHVKYLDLTLNNRLDFDNASALTELSDLEV
LDLDSYNSHYFRIAGVTHHLEFIQNFTNLKVNLSHNNIYTLDKYNLESKSLVELVFSGNRL
DILWNDDDRYVISIFKGALKNLTRLDLSLNRLKHPNEAFLNL PASLTELHINDNMLKFFNWT
LLQQFPRLLELLDLRGNKLLFLTDSLSDFTSSLRTLLLSHNRISHLPSGFLSEVSSLKHLDSL
SNLLKTINKSALETKTTKLSMLEHGNPFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFFFITTVMVLAALAHHLFYWDVWFYINVCLAKVK
GYRSLSTSQTFYDAYISYDTKDAVTDWVINELRYHLEESRDKNVLLCLEERDWDPLGAIID
NLMQSINQSKKTVFVLTKKYAKSWNFKTAFYLALQRLMDENMDVIIFILLEPVQLHSQYLR
RQRICKSSILQWPDPNPKAEGLFWQTLRNVVLTENDSRYNNMYVDSIKQY

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 826-848

FIGURE 212

CCAGGTCCAACTGCACCTCGGTTATCGATTGAATTCCCCGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACCGCTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCCTCAGGAGGTCCTCCAGGGGCCAGTGGCCTGAGGCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCGGG
TCCAGCAGCATCAGCAGCCCCCAGGACGGGGAGGCACAGGTGGCCCCACCACCCGGAGGA
GCAGCTCTGCCCCGTCCGGGGGATGACTGATTCTCTCCGCCAGGCCACCCAGAGGAGAA
GGCACCCCCCTGGAGGCACAGGCCATGAGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGTGGCAGTGGGCCACAGAGCACGCCAACGGCCGGCGTAGGGTGTGCTGT
CCGGCTCACGGGACCCCTGCTCCAGTCGTTCTGCAAGCGTGTGTACAGCCCTCTCA
CCACCTGCGACGGCACCGGCCCTGCAAGCACCTACCGAACCATCTATAGGACGCCACCGC
CGCACCCCTGGCTGGCCCTGCCAGGCCCTCGCTACGGCTGCTGCCCCGGCTGGAAAGGGAC
CAGCGGGCTCTGGGGCTGTGGAGCAGCAATATGCCAGCCCATGCCGAACGGAGGGA
GCTGTGTCCAGCCTGGCCCTGCCCTGCCAGATGGGGGTGACACTTGCAGTC
GATGTGGATGAATGCACTGCTAGGAGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCG
CAGTTACTGGTCCAGTGTGGAGGGCACAGCCTGTCAGACGGTACACTCTGTGTG
CCAAGGGAGGGCCCCCAGGGTGGCCCCAACCGACAGGAGTGGACAGTGCAATGAAGGAA
GAAGTCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGGC
CCCACTGCAAGCCTGGCTCGCAGGCACTGGAGCATGGCTCCGGACCCGGCACCTCC
TGGTCACTCTTCCAGCAGCTGGCCATCGACTCCCTGAGCGAGCAGATTCTTCTG
GAGGAGCAGCTGGGTCTGCTCTGCAAGAAAGACTCGTGTACTGCCAGCGCCCCAGGCTG
GACTGAGCCCCCTACGCCGCCCTGCAGCCCCATGCCCTGCCAACATGCTGGGGTCCAG
AAGCCACCTGGGGTGACTIONGGAGGCCAGGCAGGGCCTCTCTCTTCCCTCCCC
TTCTCGGGAGGCTCCCGAACCTGGCATGGATGGGCTGGGATCTCTCTGTGAATCCAC
CCCTGGCTACCCCCACCCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGCCCTCAGCTG
AGGGAAAGGTAGAGCTCCCTGCTGGAGGCTGGGACCCATGGCACAGGCCAGGCAGGGAG
GCTGGGTGGGGCTCAGTGGGGCTGCTGCCCTGACCCCCAGACAATAAAATGAAACGTGA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCCGCACTTAGAGT
CGACCTGCAAGCTGGCCGCATGGCCAACCTGTTATTGCACTTATAATGGTTACAAAT

FIGURE 213

MRGSQEVLMLWLLVLAvggTEHAYRPGRRVCAVRAHGDGVSESFVQRVYQPFLTTCDGHRAc
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRtSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSAADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKD

Signal sequence:

1-19

FIGURE 214

GCCAGGGCAGGTGGGCCTCAGGAGGTGCCCCAGGGCGGCCAGTGGGCCTGAGGCCCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGAACAGCAGGGCCACCATGCCACGCCCTGGCTC
AGCAGCATCAGCAGGCCCTGTGGTGGCAGCAAAGTTCACGCTTGCTGGCTGGCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTCGGGTGTCCGAGGGCTGAGGTCTCTCATCTTCTCCCTAGC
AGTGGATGAGCAACCCAACGGGGCCCGGGAGGGAACTGGCCCGAGGGAGAGGAACCCC
AAAGGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACGGGGAGG
CACAGGTGGCCCCCACCACCGGAGGAGCAGCTCTGCCCTGTCCGGGGATGACTGATT
TCCTCCGCCAGGCCACCCAGGGAGAAGGCCACCCGCCCTGGAGGCACAGGCCATGAGGGC
TCTCAGGAGGTGCTGCTGTAGTGGCTTCTGGTGTGGCAGTGGCGGCACAGAGCACGCC
CCGGCCGGGGCTAGGGTGTGCTGCTGGGCTCACGGGACCCCTGTCCGAGTCGTTCG
TGCAGCGTGTGTACAGCCTCCTCACCACTGCCACGGGACCCGGCTGCAGCACCTAC
CGAACCATCTATAGGACGCCCTACGCCGCAGCCCTGGCTGGCCCTGCCAGGCC
CGCGTGTGCCCGGCTGGAAGAGGACAGCGGGCTTCTGGGCTGTGGAGCAGCAATAT
GCCAGGCCCATGCCGAACGGAGGGAGCTGTGTCCAGCCTGCCGCTGCCCTGCA
GGATGGCGGGGTGACACTTCCAGTCAGATGTGGATGAATGCCAGTGTCTAGGAGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGAGCTTACTGGTGCAGTGTGGAGGGCACAGCC
TGTCTGCAGCGGTACACTCTGTGTGCCAAGGGAGGGCCCCCAGGGTGGCCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGGTGAGAGGCTGAGTCCAGGGTGGACCTGCT
GGAGGAGAAGCTGAGCTGGTGTGGCCACTGACAGCCTGGCTCGCAGGCACTGGAGC
ATGGGCTCCCGGACCCGGAGCCTCTGGTGCACTCTTCCAGCAGCTGCCGCTGCAC
TCCCTGAGCGAGCAGATTCTCTGGAGGAGCAGCTGGGGTCTGCTCTGCAAGAAAAGA
CTCGTGACTGCCAGCGCTCCAGGCTGGACTGAGCCCTCACGCCGCCCTGCAGCCCC
CCCCCTGCCAACATGCTGGGGTCCAGAAGGCCACCTCGGGTGA
CTGAGCGGAAGGCCAGG
AGGGCCTTCTCCTCTTCCCTCCCTGGAGGCTCCCAAGACCTGGCATGGGAT
GGGCTGGGATCTCTCTGTGAATCCACCCCTGGCTACCCCAACCTGGCTACCCCAACGGCA
TCCAAGGCAGGTGGACCCCTCAGTGAGGGAGGTACGAGCTCCCTGCTGGAGCCTGGAC
CCATGGCACAGGCCAGGAGCCGGAGGCTGGTGGGGCTCAGTGGGGCTGCTGCC
CCCCAGCACAATAAAATGAAACGTG

FIGURE 215

MRGSQEVLMLWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSES FVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACC PGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSVDCECSARRGGCPQRCINTAGSYWCQCWEGHSLSAADGTL CVPKGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 216

CCACACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA
GGTGGGCCTCAGGAGGTGCTCCAGGCGGCCAGTGGGCTGAGGCCCCAGCAAGGGCTAGGG
TCCATCTCAGTCCCAGAACAGCAGCGGCCACCATGCCACGCCCTGGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGAGGCACAGGTGGCCCCACCAACCCGGAGGAGCAGCTCCG
CCCTGTCCGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCG
CTGGAGGGCACAGGCCATGAGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGC
AGTGGGCGGCACAGAGCACGCCAACCGGCCGGCGTAGGGTGTGCTGTCGGCTCACG
GGGACCCCTGTCCTCGAGTCGTTCTGAGCGTGTGTAACAGCCCTCCCTCACACCTGCGAC
GGGCACCGGGCCTGAGCACCTACGAACCATCTATAAGGACGCCAACGCCGAGCCCTGG
GCTGGCCCTGCCCAGGCTCGTACCGCGTGTGCCCCGGCTGGAAGAGGACAGCGGGCTTC
CTGGGGCTCTGGAGCACAAATATGCCAGGCCCATGCCGAACGGAGGGAGCTGTGTC
CCTGGCCGCTGCCGCTGCCCTGAGGATGGCGGGTGACACTTGCACTGAGCTGGATGA
ATGCACTGCTAGGAGGGCGGCTGCCCCAGCGCTGCGTCAACACCGCCGCAGTTACTGGT
GCCAGTGTGGAGGGCACAGCCTGTCAGCAGCGTACACTCTGTGTGCCAAGGGAGGG
CCCCCCAGGGTGGCCCCAACCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG
GCTGCAGTCAGGGTGGACTGCTGGAGGAGAAGCTGCACTGGTGTGCCCCACTGCA
GCCTGGCCCTGCAGGCAGTGGAGCATGGCTCCCGAACCCGGCAGCTCCTGGTGCAC
TTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTCTCCTGGAGGAGCAGCT
GGGGTCTGTCCTGCAAGAAAGACTCGTGA
CTGCACTGCCAGGCCAACATGCTGGGGTCCAGAAGGCCAC
TCACGCCGCCCTGCAAGCCCCCATGCCAACATGCTGGGGTCCAGAAGGCCAC
GGGTGACTGAGCGGAAGGCCAGGCAGGGCTCTCTCTCCCTGGGAG
GCTCCCCAGACCCCTGGCATGGATGGGCTGGGATCTTCTGTGA
ATCCACCCCTGGCTAC
CCCACCCCTGGCTACCCAAACGGCATCCAAAGGCCAGGTGGCCCTCAGCTGAGGGAGGTAC
GAGCTCCCTGCTGGAGCTGGGACCCATGGCACAGGCCAGGCAGCCGGAGGCTGGGTGGGG
CCTCAGTGGGGCTGCTGCCATGCCAAC
AATGAAACGT

FIGURE 217

MRGSQEVL LMWLLVLA VGGTEHAYRPGRVCAVRAHGD PVSE SFVQRV YQPFLIT CDGH RAC
STYRTIYRTAYR RSPGLAPARPRYACC PGWKR TSGLPGAC GAAICQPPCRN GGSCVQ PGRCR
CPAGWRGDT CQSDVDEC SARRGGC PQRCVN TAGSYWCQCWE GHSL SADG TL CVPKG GPPRVA
PNPTGVDSAMKEEVQRLQSRV D LLEEK LQLV LAPL HSL ASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKD S

Signal sequence:

1-19

FIGURE 218

GGTTGCCACAGCTGTTAGGGCCCCACCCTGGGCCCTTGTCAAGGAGAGACAGCCTCCGGCCGGAG
GACAAGTCGTGCCCACCTTGGCTCGCGACGTGATTCCTGGACGGTCCGTTCTCGCTCAGCTCCGGCG
AGTTGGGTCCTCCGGCTCCCGCTCCCTGGTCTCCCTTCCTGGCTGGCCCGGTTATGGGAGG
AGATTGTCCTCCAGGGCTAGCAATTGGACTTTGATGATGTTGACCCAGCGCAGGAATAGCAGGCAACGTGAT
TTCAAAGCTGGCTCACGCCCTGTTCTCTCGTGTAACTGCAAAACCATTGGAGCAGGAATTCAACAT
~~TGTCCTGATGGTGGTGAAGAAAGGGTGAACACGGAATATGGGAGAACA~~CTTTCGCTGTG
ATGACGGCGTCATGATGCGCCGGCAAAAGGGCATTTCTACCTGACCCCTTCTCCTCATCTGGGACATGTAAC
TCTCTCTGGCTTGTGAGTCGGCTACCTGGCTCAGCTGCTCTCGCTCAGTGACCCCTGGAGTGATCTCGCATGCTCT
TCTCTTCTCCTAGGCTACACTGGTGAAGGACAGCTTCAGTGACCCCTGGAGTGATCTCGCATGCTCT
AAGCAGCTTCATAGAAATGGAGATAAGACTACCAATTGGTGCCTGGCCAGGGCAGCGACACCCTCGTA
TCAAGAATTTCAGATAAACACCCAGATTGTGAACACTGAAATACTGTTACATGCAAGATCTCCGGCCTCC
GGGGCTCCATTGCAACATGTGACAACCTGTGGAGCGCTTGCACCATCAGTCCCTGGGTGGGAATTGTG
TTGGAAAGGGAACTACCGCTACTTCTACCTCTTCATCCTTCTCTCCCTCCTCACAAATCTATGTCCTCGCCT
TCAACATCGCTATGTTGGCCCTAAATCTTGAAAAATGGCTCTTGAGACMTGAAAGRAACTCTCGGAAC
TTCTAGAAGTCCTCAITTGCTCTTACACTCTGGTCCGTCGTGGGACTGACTGGATTCTACACTTCTCGTGG
CTCTCAACCCAGAACAACTGAGACATCAAGGATCATGGACAGGAAATCTGGCTCGAAGATCCCTACAGCC
ATGGCAATATACTTGAGAAACTGCTGTAAGCTGGCTTGGCCCTTGGCCCTTGGAGATCTGGCTGGAGGTA
TTTGCCACTTGAGGAAGTGGAGACTGGCCACTTCCCTAGTCAAGGACAGTAGTACAGCCTTGGCAACAGGCC
CAGCCCCCTCAAGAACACCTGAACTCAATGAGATGGCGAGGAGCACGACACTCCGGAGAGATGCACTCC
AGCCCCCAGAGGCCACACAGGGAGCTGAAGCTGAGAAGTACGCTATCTATGAAAGAGACTTTGGTT
TAATTAGGGCTATGAGAGATTGGTGAAGATTAACCTGAGACAGGAGCAAGTAAGCTGTCCTTTAACT
GTTTTCTTGTGCTTTAGTCACCGCATGTCACACTGGCATTTCTCTGCTCAAGCTTTAAATTCTGA
CAAGGAGTGGAGAAGTGTCACTCCTCTGTAACTGGAAAATGGGAAATCTGGCTCTGGCCCTTGCA
CTCTGGCCCTCAGGCAAGGGTCCCTTGGAGGCCCTCTCTCCCTCAGATCTGGCCCTCTGCTTGGGTCA
TGGTCTCAITCTGGGCTAAAGATTGGAGACTGGCTAAATCTCCCAAGCTGCTCACGTGCTGACTCCAGA
GGGAGTCACAGAGACGACTCTGGCAGGGGATCTAACCTGGGTTCTGGGCTCTCAGGACTGAAGAGGAGGAG
TGGGGTCAGAAGATTCTCTGGGCCAAACTGCGACATGCCCCAAATCTTGAATGGGACAGTACCT
TCCACTTGTGTGTAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGTTTTCTCTTGTACTCTGCTCTCC
CCATGGCCCTCAGGCAAGGGTCCCTTGGAGGCCCTCTCTCCCTCAGATCTGGCCCTCTGCTTGGGTCA
TGGGAGTGGAGGAAAGCCAGACTGCTGGCTCAGGACTCTGGCTAACCTGGGACTGGCTCTGGCT
ACTATCCCTCAACTCCCTGTGAGGCTGAGGAGGCCCTGAACTGACAAATGGGAACCAAGGCAAGAGAG
GCTCTCTCTCTCTCCCTCAGGACTGGCTGAGGAGGCCCTGAACTGACAAATGGGAACCAAGGCAAGAGAG
GGGAGTGGAGGAAAGCCAGACTGCTGGCTCAGGACTCTGGCTAACCTGGGACTGGCTCTGGCT
AIGGTAACCCACTGGGGCTCTCTCCAGGAGGCCCTCTCCAGGACTTCCACGGGAGACTCCAGGACACTT
CACCCTGGGGTGGCTGTGCGCCCAAGTCAAGGACTCTGGCTAACCTGGGACTGGCTCTGGCT
ATTATATGTGGCTATATTCTCTAGAGCACCTGTGTTCTCTCTATTCTAAGCCAGGGCTGTCTGGATGACTTAT
GGGGTGGGGAGTGTAAACCGGAACCTTTCATCTATTGAGGGGATTAACCTGTGCTAATGCA

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTLFLILGTCTLFFAFECRYLAV
QLSPAIPVFAAMLFIFSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGQRPPP
RIKNPQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCPWVGNCVGKRNYRYFYL
FILSLSLLTIYVFAFNIVYVALKSLKIGFLETLKETPGTVLEVЛИCFFTLWSVVGLTGFHTF
LVALNQTTNEDIKGSWTGKNRVQNPYSHGNIVKNCEVLCGPLPPSVLDRRGILPLEESGSR
PPSTQETSSSLPQSPAPTEHLNSNEMPEDSSTPEEMPPPEPPQEAEEAEK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 220

AAAACCTGTATTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTGAAATTAGGTAT
TATAGGGATGGTGGGGTTGATTTNTCTGGAGGCTTTGGCTTGACTCTCNCTTCT
CCCACAGAGCNCTTCGACCACACTGCCCTGGTGGGAATTGTGTTGGAAAGAGGAACTA
CCGCTANTTCTACCTCTCATCCTTNTCTCTCCNCCTCACAACTATGTCTTCGCCTCA
ACATCGT

FIGURE 221

GTTGTGCTTCAGCAAAACAGTGGATTAAATCTCCTTGACAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAGAAAAGAAAAACCGAACCTGACAAAAAAGAAGAAAAGAAGAAGA
AAAAAAATCATGAAAACCATCCAGCAAAATGCACAAATTCTATCTTGGCAATCTTCAC
GGGGCTGGCTGCTCTGTGCTCTTCCAAGGAGTGCCCCGTGCGCAGCGGAGATGCCACCTTC
CCAAAGCTATGGACAACGTCAGGGTCCGGCAGGGGGAGAGCGCCACCCCTCAGGTGCACTATT
GACAACCGGGTCAACCGGGTGGCCTGGCTAAACCGCAGCACCATCTCTATGCTGGAAATGA
CAAGTGGTGCTGGATCTCGCTGGTCTTCTGAGAACACCAAACGAGTACAGCATCG
AGATCCAGAACGTTGATGTTATGACGAGGGCCCTAACCTGCTCGGTGAGACAGAAC
CACCCAAAGACCTCTAGGGTCACCTCATTGTCAGGTATCTCCAAAATTGTAAGAGATTTC
TTCAGATATCTCCATTAAATGAAAGGGAAACATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGACCTACGGTTACTTGGAGACACATCTCTCCAAAGCGGTTGGCTTGTGAGTGAAGAC
GAATACTTGGAAATTCAAGGCATCACCGGGAGCAGTCAGGGACTACGAGTGCACTGCCTC
CAATGACGTGGCGCGCCGTGGTACGGAGAGTAAAGGTACCGTGAACATCCACCATACA
TTTCAGAAGCCAAGGGTACAGGTGCCCCGTGGACAAAAGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCTCAGCAGAACTTCAAGGTACAAGGATGACAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAACAGACCTTCTCTCAAACACTCATCTTCAATGTCCTG
AACATGACTATGGAAACTACACTTGCCTGGCCTCCAAAGCTGGGCCACCCAATGCCAGC
ATCATGCTATTGGTCCAGGCCTCGAGCAGGTGAGCAACGGCACGTGAGGGAGGGCAGG
CTCGCTCTGGCTGCTGCCCTCTGGCTTGACCTGCTCTCAAATTTGATGTGAGTGCC
ACTTCCCCACCGGGAAAGGCTGCCACCCACCAACACAAAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATAACAAATGAAATTAGAAGAAACACGCCATGGGACAGA
AATTGAGGGAGGGAAACAAAAGAATCTTGGGGAAAAGAGTTAAAAAGAAATTGAA
AATTGCCCTGCAGATATTTAGGTACAATGGAGTTCTTCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTCAACCTTGGTGCAGTGCGCAA
GGGCTCAGCCTCTGCCACAGAGTGCCACCGTGAACATTGGAGCTGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCAAGCGTGGCGCTGCGG
GCACTTGGTAGACTGTGCCACCAGGGCTGTGTTGTGAAACGTGAAATAAAAGAGCAAAA
AAAAA

FIGURE 222

MKTIQPKMHNSISWAIFTGLAALCLFQGVPVRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCCSVQTDNHPK
TSRVHLIVQVSPKIVEISSLISINEGNNISLTCIATGRPEPTVTWRHISPKAvgFVSEDEYL
EIQGITREQSGDYECASNDVAAPVVRRVKTVNYPPYISEAKGTGVPVGQKGTLQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNYTCVASNKLGHHTNASIML
FGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF

Signal peptide:

amino acids 1-28

FIGURE 223

GAAAAAAAATCATGAAAACCATCCAGCCAAAATGCACAATTCTATCTTGGGCAATCTTC
ACGGGGCTGGCTGCTCTGTCTTCCAAGGAGTGCCCGTGCAGCGGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGAGAGCGCCACCTCAGGTGCACTA
TTGACAAACGGGTACCCGGTGGCCTGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCCTGGATCCTCGCGTGGTCCTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTAACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCAAAATTGTAGAGATT
TCTTCAGATATCTCCATTAATGAAGGGAACATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAAGAG

FIGURE 224

ATGGCTGGTGA CGGGCGGGCGGCAGGGGACCGGGGCCGGCCGGCCCCGGAGGGGCCAGCTGCCGGAGGCC
ATCACCGCCTGGCCGACTCACCAGTAAGAAGCTCGCGCTGCAGAGCTGGAGCTGGAGCAACGTGGATTCCAG
AAAGGGACACAGACAGCTGGTACGGCAGCAGCTGGAGCTGGTCTTAGCAGGTGGCTCTACTGTGCGCT
GCAGCTCTCTGGCTCTGGCCCTAGGGGTCAAGTACACACAGAACCCATCCACAGCACCTTGACCTTACA
GAGGCTGCTGATCGAGCTGGTGA AAAAATCTGGAGTCCCTGGACAGGGGTGAGCCCCCTGTGAGGACTTTAC
CACTTCCTCTGGGGCTGGATCTGGAGCTGGCCCTACAGGTGGAGCATTGAGGAGCTGGAGCCAGCAGTGGAG
GACCTCATGGAGAAGATTCTGGTTGGAAACATTACGGGGCCCTGGGAGCAGAACCTTTATGGAGGTGTTGAG
GCAGTGGAGGACCTAACGGGACCCATTCTTACACGCTCACATCAGTGGCAGCTCAAGAGTCCAAACAGC
AATGTTATCCAGGGACAGTCTGGGCTCTTCTGGGATTAACAGAGAACCTGGGCACTGGCTGGGGCCCACTCCACAGG
AAAGTGTCTACIGCTTACATGGGATTAACAGAGAACCTGGGATGCTGGCTGGGGCCCACTCCACAGG
GAGCAGATCCAGCAGGTGGAGATCTGGAGATCACAGTGGCCCAACATCACAGTGGCCAGGACAGCAGCGCGAC
GAGGAGAAGATCTACACAGATGAGATTCTGGAGCTGGAGCTTCAGTGGCCCTTCATGGACTCTGGCTTGAGTTC
CTGGTCTCTCTGGTACCAAACTGGAGTTGAGTGTGAGCTCTGGAGCTGGCTGGTGTATGGATGGATTATTGAG
CAGGTGTAGAGCTCATACCCAGCAGCAGCATCTGGAGCTGGAGCAACATTCTGGAGCTGGCTGGTGTCAAAAG
ACAACCTCAAGGGCTGGACGACCTCTTGGAGCTGGAGCAACACGGATGGCCCTGGAGCTGGAGACCCCTATGGACTAAAGAG
TCTCTGGTGGAGGAGCTGGAGCAGCTTCTGGAGCTGGAGCTGGCTGGCTGGAGCTGGAGCTGGAGCTGGAGCT
GTGGAGGAGCACCTTGGAGGCAACAGAAAATCTGGAGGGGATGATCAGCAGAACTTGGAGGCACTTGGAG
GAGGGGGAGCAGCTGGAGGAGCTGGGGTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGG
GATATGATGGTTTCCCAGACTTCTGGAGCCAAAGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG
GAAGATTCTCTTCCAAAACAGCTGGATGATGATGACTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG
GCTGGCTTACAAAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG
GCTGGCTTACAAAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG
AACCACCAAGCTCTCTGGGGATTTGGCCAGTGGGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG
CTGGTGGAGCCAGGACCCCCAACAGCCCTGGGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG
CACTTGGCTGGCCCTGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG
GAAATGGCCAGCTGTGACCAAGACTGGGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG
ATGCAAGCTGGCTGGGCTAGTCCCTCCCCCAGGGTGAAGTACAGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG
TGGCTCTGGCTTGGGGTGGCCCTGGCTCAGCAGGGCCACCTTACAGTGGAGCTGGAGCTGGAGCTGGAGCTGGAG
GCGCTGGAGAGGCTGGGGAGGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTOLEVLAGASLLAALLLGCLVALGVQYHRDPSH
STCLTEACIRVAGKILESLDRGVSPCEDFYQFSCGGWIRRNPFDGRSRWNTFNSLWDQNQA
ILKHLLNETTFNNSSEAEQKTQRFYLSQLQVERIEELGAQPLRDLIEKIGGWNTITGPWDQDN
FMEVLKAVAGTYRATPFFTYYISADSKSSNSNVIQVDQSGFLFLPSRDYYLNRTANEKVLTAY
LDYMEELGMILLGGRPTSTREMQQVLEIQLANITVPQDQRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPLLELSDSEPVVVYGMMDYLQQVSELINRTEPSIILNNYLIWNVLVQKTTSSL
DRRFESAQEKLLLETLYGTTKSCVPRWQT CISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTAFFEEALGQLVWMDEKTRQAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF
FQNMLNLYNFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLPTKNEIVFPAGILQAPFYARNH
PKALNFGGITGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRLKHGEQQLPAVGLTNHQLFFVGAQVWCS
VRTPESSHEGLVTDPHSPARFRVLGTLNSNRDFLRHFGCPVGSPMNPQQLCEVV

Type II Transmembrane domain:
amino acids 32-57

FIGURE 226

FIGURE 227

GGCGGAGCGGGGGTGCAGCGCGCCGTGATGGCTGGTGACGGCGGGCGGGCAGGGGA
CCGGGGCGCGGCCGGAGCGGCCAGCTGCCGGAGCCCTGAATCACCGCCTGCCGAC
TCCACCATGAACGTCGCGCTGCAGGAGCTGGAGCTGGCAGAACGTGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACCGAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCAC
TGCTGGCTGCACTGCTCTGGCTGCCTGTGGCCCTAGGGTCCAGTACACAGAGACCCA
TCCCACAGCACCTGCCTTAACAGAGGCCCTGCATTGAGTGGCTGGAAAAAAATCCTGGAGTCCT
GGACCGAGGGGTGAGCCCCCTGTGAGGAACTTTACCACTGTTCTCTGTGGGGCTGGATTGGGA
GGAACCCCCCTGCCCGATGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAC
CAGGCCATACTGAAGCACCTGCTGAAAACACCACCTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTCTACCTATCTGCCTACAGGTGGAGCGATTGAGGAGCTGGAGCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGCCCTGGACCAG
GACAACTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCATTCTTCAC
CGTCTACATCAGTGCCTACTAAGAGTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG
GGCTCTTCGCCCCCTCGGGATTACTACTAAACAGAACTGCCAATGAGAAAGTAAGGAAC
ATCTTCGAACCCCCATCCCTACCCCTGGCTGAGCTGGCTGATCCCTGGTGACTIONTCCCT
TTGCCAAGGGTCAGAGCAGGGAAAGGTGAGCCTATCCTGTACCTAGTGAACAAACTGCCCT
CCTTTCTTCTTCTTCTTCTCCCTCCCTCCCTTCTTCCCTTTCTTCCCTTCCCTTCC
TCTTATTCTCTAGTAGGTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGAAATCG
GAGATATAAGTTCCGAGCCATTGCCACAGGAAGCGTCAGTGTGATGGGTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAGAGGGCTGTGAGGATTCAAGGAGAGACTTATGGAGCC
AGCAAAGTCTCCTGAAGAGATTGCAATTGAGCCAGGTCTGTAG

FIGURE 228

ATGCCTACTACCTTCAACTAAGAACATGAGATCGTCTCCCCGTGGCATCCTGCAGGCC
TTCTATGCCCGCAACCACCCCAAGGCCCTGAACCTCGTGGCATCGGTGTTGATGGCCA
TGAGTTGACGCATGCCCTTGATGACCAAGGGCGCGAGTATGACAAGAAGGAACCTGCGC
CCTGGTGGCAGAATGAGTCCCTGGCAGGCCCTCGGAACCACACGGCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGAGAGGCTCAACGGCGCAGACGCTGGGGAGAACAT
TGCTGACAAACGGGGGCTGAAGGCTGCCATAATGTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGGAGCAGCACTGCCAGCCGTGGGCTACCAACCACAGCTCTTCCTGGGATT
GCCAGGTGGTGGCTGGTCCGCACACAGAGAGCTCTACAGGGGCTGGTACCGACCC
CCACAGGCCCTGCCGCTTCCGCGTGGGACTCTCTCCAACTCCGTGACTCTCGCGC
ACTTCGGCTGCCCTGCGCTCCCCATGAACCCAGGGCAGCTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGCCAGCTGTACCCAGACCTGGGCACTCTCTGACAAAGCTGT
TTGCTCTGGGTTGGGAGGAAGCAAATGCAAGCTGGCTGGGCTAGTCCCTCCCCCACA
GGTGACATGAGTACAGACCCCTCTCAATCACCAATTGCGCTGCTTGGGGTGCCT
GCCCTCCAGAGCCCCACCAATTCACTGTGACATCTTCCGTGTCACCCCTGCCCTGGAAGAG
GTCTGGTGGGAGGCCAGTCCCATAAGGAAGGAGCTGCCCTTCTGTCCCCAGGCTCACT
CAGCCTGGCCGATGGGCTGCCGTGCCCTGCCACTGTGACCCACAGGCTGGTGGTG
TACCTCCTGGACTCTCCCCAGGCTCACTCAGTGCACCTTAGGGGTGGACTCAGCTCTGTC
TGGCTCACCCCTACGGGCTACCCACCTCACCCGTGCTCCTGTGCCACTGCTCCAGTG
CTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAAGCCAAGGGCTCTGAAAGCCTCTGC
TGCCCCACTGTTCCCTGGCTGAGAGGGAAAGTGCATATGTTAGCGGGTACTGGTCCCTG
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTATTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCT
GGCCCTTATAGGACC

FIGURE 229

CCCCACGCGTCCGAGCCGCCCGAGAAATTAGACACACTCCGGACGGGCCAAAGCAACCGAGA
GGAGGGGAGGCCAAAACACCGAAAAACAAAAGAGAGAAACAACCCAACAACGGGTGG
GGGGAAAGAAAAGAAAAGAACCCACCCACCCACCAAAAAAAAAAAAAAAA
AAAAAAAAAAATCTGGCCGCCGCTGGTTCCCGGGAGACTCGCCAGCACCAGGGG
TGGGGGACTGCAGACTGAACCTCTGGAGAGTGGACAGCCCCTAGCAGGGATGGACATGATG
CTGTTGGTGCAAGGGTCTGTTCTGCAACCGAGTCGGCTGGCCTCTCAGCCCTGTG
CTGCCCTGTAACCCCTCTGCCCTCCCGCTGGACAGAGTGTGACTCCCTGGCGCCGTGG
ACAACATGATGGTCAGAAAAGGGACACCCGGTCTTAGGGTATTGGAGATGGAGCT
TCAAAGGGTGCCTGCTGAACCCGTCAAGTATTATTTGGGGAGGTGATAAGTGGTCAGT
GGATCCTCGAGTTCAATTCAACATTGAATAAAAGGGACTCACAGCTCAGACATAAGATG
TAGATGTGACAGATGATGCCCATACACGTGTTCTGTCAGACTCAACATACCCAGAAC
ATGCAGGTGCATCTAACGTGCAAGTCTCTCAAGATAATGACATCTCAAATGATGAC
CGTCAATGAAGAACCCACGTCACTCTTACTGTTGGCCACTGGAAACAGAGCTTCCA
TTTCTTGCGACACATCCCCATCAGCAAAACCATTGAAAATGACAATATTGGACATT
TATGGAAATTACAAGGACAGGGTGGGAATATGAATGCACTGGGAAATGCTGTGCTT
CCCAAGATGTGAGGAAAGTAAAGGTGTTCTCAATTGGCTCTACTATTCAAGGAAATT
CTGGCACCGTGACCCCCGGACCACTGGCTGATAAGATGTGAAGGTGCAAGTGCGCCCT
CCAGGCTTGAATGGTCAAAGGAGAGAAGAGCTTCAATGCCAACAGGAATTATT
TCAAAATTCTAGCACAAAGATCCACTCTACTGTTAACACGGGACACTGGGAACTGGCA
ATTATACCTGTGTCGCTGCCAACAGCTGGCACAAACCAATGCCAGCCTGCTTAAACCC
CCAAGTACAGCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTCTCTGCTGGTACCT
TGTGTTGACACTGCTCTTCCACAGCATATTCTACTGAAAGATGCCATTCTAACTAA
TTCAAGACCCATAAAAAGGCTTTAAAGGATTCTCTGAAAGTGTGATGGATCCAATCT
GGTACAGTTGTTAAAGCAGCGTGGGATATAATCAGCAGTGCTTACATGGGATGATGCC
TTCTGTGAAATTGCTCATTATGTAATCTTAAATTCTACTCTTTGATTAGCTACATTA
CCTTGTAAGCAGTACACATTGCTCTTAAAGAGCTGAAAGCTGAAATTACTTTAG
AGGATTAATTATGTTGATTCTGTTGATTAATCTCACACTTTCAAAGCATTGCTAGT
CTGCTAGGTTGCAAGCTGTAGTTACAAAAAGAATTGCACTGAATATGTGATTCTTAA
GGCTGCAATACAGCATTGCTCTCTTCAATAAGAGTCAATCACCATTACAAAGATG
CAATTTTCTCTTTGATAAAAAAGCAAAATAATTGCTCTGAGATTATTCTCAAATAA
TAACACATATCTAGATTCTCTGCTTGATGATATTCACTGTTCAAGGATGAGCTGTAAT
ATACTGGCTGTCAGCTGCTCTCTTCTGTAAGTTGAGCTGAGGGTGTGCCCTCATAC
ATAATTATTCTCTTGTCTCCAACTATAATAAAAATGTTGCTAAATTCTAACATTG
AAAGTAAATAAACAGACTGATCAAGTTAAACCATACACTATCTCAAGTAACGAAGGAGC
TATTGAGCTGAAAAATCTCTCTGCACTGACAATGGGTTGAGAATTGGCCCAACT
AACTCAGTTCTGTGAGAGACAAATTAAACAGTATGAAATATACCATATGATTTC
TTTACTGTGAGCTAAATGTTGAGTCAACCGTGGGAAATCTCCCTTAAATGACAGCACA
GTCCACTCAAAGGATTGCTAGCAATACAGCATCTTCTCTTCACTAGTCCAAGGAAAAA
TTTAAGATGATTGTCAGAAAGGGCACAAAGTCTATCACCTAAATTACAAGAGTTGGTA
AGCCCTCATCTAAATTCTTGTGGCAGGTATTAGACAGTCGACCTGGAGGGTATGGA
TATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACAGGGTGGTTATGACCGTACTC
AGGAGGAAATTACAGAGACAAATTGACAAACTGAAATGAGACATGCACATAATAGATACA
CAAGGAATAATTCTGATCCAGGATCGCCTTCCAAATGGCTGATTATAAAGGTTTTGG
AGCTGCACTGAAGCAGTCTTATTTATGATATACACCTTTGTTTAAATTGACCTGCCA
AGGTAGCTGAAGACCTTTAGACAGTTCATCTTTTTAAATTCTGCTTAATTAA
AGCACAAATTATGGGAGCTTGTCAAAAAAAAAAAAAAAA

FIGURE 230

MMLLVQGACCSNQWLAAVLLSLCCLLPSCLPAGQSVDPPWAADVNMVRKGDTAVLRCYLED
GASKGAWLNRSSIIIFAGGDWKWSVDPRTVSISTLNKRDYSLQIQNVDTVDDGPYTCCSVQTQHTP
RTMQVHLTVQVPPKIYDISNDMTVNEGTNVTLTCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRQAGEYECSAENAVSFDPVRKVKKVVNFAPTIQEIKSGTVTPGRSGLIRCEGAGV
PPPFAFEWYKGEKKLFNGQQGIIIQNFSTRSILTVTNVTQEHFGNYTCVAANKLGTNASLPL
NPPSTAQYGITGSADVLFSCWYLVLTLSFTSIFYLKNAIQL

Important features of the protein:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.

amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325

Myelin P0 protein:

amino acids 92-121

TOP TO BOTTOM 7/6/01

FIGURE 231

AGTGGTTCGATGGGAAGGATCTTCTCAAGTGGTCTCTTGAGGGGAGCATTCTGCTGG
CTCCAGGACTTGGCCATCTATAAGCTTGGCAATGAGAATAAGAAAATTCTCAAGGAGGA
CGAGCTCTTGAGTGA GACCCAACAAGCTGTTTCAACCAATTGCAATGGAGCCTTCGAA
TCAATGTTCAAAGGCCAAGAGGAATGGGTGA ACTTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCTGCTCACCGCTGGCGTGGCTGCTGGTGGTCAAGTCTGAATCTGCAGGC
GCGGCTCGGGTCTGGAGATGTATTCTCAATGACACTCTGGCGCTGAGGACAGCCCCTG
CCTTCTCTGCTGAGTCAGCACACCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTG
CAAGTCCTGCAAGGCCAACACTCACCTGGGTCGCGTCAGCCATGAGCACTTGTGCA
AGCGGGT
AGACAACCTCACTCAGAACCCAGGGATGTTAGAATCAAAGGTGAACAAGGCGCCCCAGGTC
TTCAAGGTACAAGGGGCCATGGCATGCTGGTGCCTGGCCCTGGCCGCCGGGACCACCTGCT
GAGAAGGGAGCCAAGGGGCTATGGGACAGATGGAGCAACAGGCCCTCGGGACCCAAAGG
CCCACCGGGAGTCAGGGAGAGGCGGGCTCCAAGGACCCAGGGTGTCCAGGGAAAGCAAG
GAGCCACTGGCACCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGGATGGGGTCTCATT
GGCCCAAAGGGGAAACTGGAACTAAGGGAGAGAACAGGAGACCTGGTCTCCAGGAAGCAA
AGGGGACAGGGGATGAAAGGAGATGCAAGGGTATGGGCCTCTGGAGCCAGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCACAGGTTTGGCTGGTTTCTGGAGCTAAAGGAGAT
CAAGGACAACCTGGACTGCAAGGTGTTGGCCCTCTGGTGCAAGTGGGACACCCAGGTGC
CAAGGGTGA GCGCTGGCAGTGCTGGCTCCCTGGCAGCAGGACTTCCAGGGAGCCCCGGGA
GTCCAGGAGCCACGGCTGAAAGGAAGCAAAGGGGACACAGGACTCAAGGACAGCAAGGA
AGAAAAGGAGAATCAGGAGTTCAGGCCCTGCAAGGTGAGGGAGAACAGGGAGGCCAGG
GCTGGCAGGTCCAAGGGAGCCCTGGACAAGCTGGCAGAAGGGAGACCAAGGGAGTGAAAG
GATCTTCTGGGAGCAAGGAGTAAAGGGAGAAAAGGGTGAAGAGGTGAAAACTCAAGTGTCC
GTCAGGATTGTCGGCAGTAGTAACCGAGGCCGGCTGAAGTTTACTACAGTGGTACCTGGGG
GACAATTGCGATGACGAGTGGAAAATTCTGATGCCATTGCTTCTGCCATGCTGGTT
ACTCAAAGGAAGGCCCTGTACAAAGTGGAGCTGGCACTGGCAGATCTGGCTGGATAAT
GTTCA GTGTCGGGAGCA CGGAGAGTACCCCTGAGCTGGAGCTGCCACCAAGAATAGCTGGG
TGACTGCAGCCACGAGGAGGACGCAAGGCCGGAGTGCAGCGTCTTGACCGGAAACCTTCA
CTTCTCTGCTCCCGAGGGTGCCTCGGGCTCATATGTGGGAAGGCAGAGGATCTGAGGAGT
TCCCTGGGACA ACTGAGCAGCCTCTGGAGAGGGGCCATTAAAGCTCAACATCATTGA

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886
><subunit 1 of 1, 520 aa, 1 stop
><MW: 52658, pI: 9.16, NX(S/T): 3
MRNKKILKEDELLSETQQAAFHQIAMEPFEINVPKPKRNGVNFSLAVVVIYLILLTAGAGL
LVVQVLNLQARLRLVLEMYFLNDTLAAEEDSPSFSLLQSAAHPGEHLAQGASRLQVLQAQLTWVR
VSHEHLLQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRD
GATGPGSGPQGPVGVKGEAGLQCPQGAPGKQGATCTPPGPPQGEKGSKGDGGLIGPKGETGTKGE
KGDLGLPGSKGDRGMKGDAVGVMGPPGAQGSKGDFGRGPPGLAGFPGAKGDQGQPCLQGVPG
PPGAVGHPGAKGEPEGSAGSPGRAGLPGSPGSPGATGLKGSKGDTGLQGQQGRKGESGVPGPA
VKGKEQGSPGLAGPKGAPGQAGKDGQVGKSSGEQGVKGEKGERGENSVSVRIVGSSNRGR
AEVYVSGTWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVGAGTGQIWLDNVQCRCGTESTLW
SCTKNSWGHHDCCSHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

C1q domain proteins.

amino acids 151-184, 301-334, 316-349

FIGURE 233

CCACACGCGTCCGAAGGCAGACAAAGGTTCAATTGTAAAGAAGCTCCTTCCAGCACCTCCTCT
CTTCTCCTTTGCCAAACTCACCCAGTGAGTGTGACCATTAAGAACATCCTCTGCCAAG
ACCAAAAGGAAAGAAGAAAAAGGGCCAAAAGCCAAAATGAAACTGATGGTACTTGTTCAC
CATTGGGCTAACTTGTGCTAGGAGTTCAAGCCATGCCCTGCAAATGCCCTCTTGCTACA
GAAAGATACTAAAAGATCACAACTGTACAACCTTCCGGAAAGGAGTAGCTGACCTGACACAG
ATTGATGTCAATGTCCAGGATCATTTCTGGGATGGGAAGGGATGTGAGATGATCTGTTACTG
CAACTTCAGCGAATTGCTCTGCTGCCAAAAGACGTTTCTTGGACCAAAGATCTCTTCG
TGATTCCCTGCAACAATCAATGAGAATTTCATGTATTCTGGAGAACACCATTCCGTGATTTC
CCACAAACTGCACTACATCAGTATACTGCATTCTAGTTCTATATAGTGCAATAGAGCAT
AGATTCTATAAATTCTTACTTGTCTAACAGAACAGTAAATCTGTGTTAACAAAGTAGTAATAAA
AGTTAACATTCAATCTAAAAAAAAAAAAAA

FIGURE 234

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758
<subunit 1 of 1, 98 aa, 1 stop
<MW: 11081, pI: 6.68, NX(S/T): 1
MKLMVLVFTIGLTLGGVQAMPANRLSCYRKILKDHNCNLPEGVADLTQIDVNQDHFWDG
KGCEMICYCNFSELLCCPKDVFVGPKISFVIPCNNQ
```

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

FIGURE 235

CCACAGCGTCCGGCAGCGTGGGCTGGACCCCAGGTCTGGAGCGAATTCCAGCCTGCAGGG
CTGATAAGCGAGGCAATTAGTAGAGATTGAGAGAGACTTTAACCCCCCGTGGTGGTGGAGGG
GCGCAGTAGAGCAGCAGCACAGGCGCGGTCCCGGGAGGCCGCTCTGCTCGCAGGAGATG
TGGAACTCTCTCAAGAAACCGACTCGGCTGTGGCACCGCGCCGCCGCGCTGGCTGTG
CGCTGGGGCGCTGGCTGGCGGCTTCTCTCGCCTTCAGCTTGCTGGTGGTTTA
TAAAATCTCCAATGAAGCTACTAACATTACTCCAACAGATAATATGAAAGCATTTGGAT
GAATTGAAAGCTGAGAACATCAAGAAGTCTTACATAATTACACAGATAACCACATTAGC
AGGAACAGAACAAAACATTCTAGGCAAGAACATTCAATCCCAGTGGAAAGAATTGGCC
TGGATTCTGTTGAGCTAGCTCATTATGATGTCCTGCTTACCCAAATAAGACTCATCCC
AAACTACATCTCAATAATTAGAAGATGAAAATGGAGATTCAACACATCATTATTGAAACC
ACCTCTCCAGGATAGAAAATGTTCCGGATATTCTACCACTTTCAGTGCTTCTCTCCTC
AAGGAATGCGAGGGCGATCTAGTGATGTTACTATGCAAGAACCTGAAAGACTTCTTTAAA
TTGGAACGGGAGCATGAAAATCAATTGCTCTGGGAAATTGTAATTGCCAGATATGGGAAAGT
TTTCAGGAAAATAGGTTAAAATGCCAGCTGGCAGGGGCAAAGGGAGTCATTCTACT
CCGAGGCTCTGACTACTTCTGCTCTGGGTGAAGTCATCTCAGACCGGTTGGAATCTCCT
GGAGGTGGTGTCCAGCTGGAAAATCTCAAATCTGAATGGTGAGGAGACCCCTCTCACACC
AGGTTACCCAGCAAATGAATATGCTTATAGGCGTGGATTGAGGAGCTTGGTCTTCCAA
GTATTCTGTTCATCCAATGGATACTATGATGACAGAAGCTCTAGAAAATGGTGGC
TCAGCACCCACAGATAGCGACTGGAGAGGAAGTCCTAAGTGGCCCTACAATGTTGACCTGG
CTTACTGAAACTTTCTACACAAAAGTCAGATGCACTCTACCAATGAGTGA
CGAGAATTACAATGTGATAGGACTCTCAGAGGAGCAGTGGAAACAGACAGATATGTCATT
CTGGAGGTACCGGGACTCATGGGTTGGTGTATTGACCTCAAGGTGGAGCAGCTGT
TGTTCATGAAATTGAGGAGCTTGGACAACTGAAAAGGGAGGGTGGAGACCTAGAAGAA
CAATTGGTTGCAAGCTGGATGCGAGAAAATTGGTCTTCTGGTTACTGAGTGGGCA
GAGGAGAATTCAAGACTCTCAAGAGGCTGGCTGGCTTATATTATGCTGACTCATCTAT
AGAAGGAAACTACACTCTGAGAGTGTGATCACCCGCTGATGTCACAGCTGGGACACAA
TAACAAAAGAGTCGAAAGCCCTGATGAGGCTTGTGAAGGAAATCTTTATGAAAGTTG
ACTAAAAAAAGTCCTCCCCAGAGTTCACTGGCATGCCAGGATAAGCAAATTGGATCTGG
AAATGATTGTTGAGGAGCTTCTCCAACGACTTGGAAATTGCTTCAAGGAGCAGCTAC
AAAATTGGGAAACAAACAAATTCAAGGGCTATCCACTGTTACAGTGTCTATGAAACAT
GAGTTGGTGGAAAAGTTTATGATCCAATGTTAAATACACCTCACTGTCAGGTTCG
AGGAGGGAGGGATGTTGAGCTAGGCAATTCCATAGTGTCTCCCTTTGATTGTCAGGATTATG
CTGTAGTTTAAGAAAGTAGCTGCAAAATCTACAGTATTCTATGAAACATCCACAGGAA
ATGAGAACATACAGTGTATCATTTGATTCATTTCTCTGAGTAAGAATTTCAGAAAT
TGCTTCCAAGTTCACTGAGGAGACTCCAGGACTTTGACAAAAGCAACCAATAGTATTAAAGAA
TGATGAAATGATCAACTCATGTTCTGAAAGAGCATTATTGATCATTAGGGTTACAGAC
AGGCCCTTTTATAGGCATGTCATCTATGCTCCAAGCAGGCCACAACAAGTATGCAAGGGAGTC
ATTCCCAGGAATTATGATGTCCTGTTGATATTGAAAGCAAAGTGACCCCTTCAAGGCC
GGGGAGGAAGTAGGAGACAGATTATGTTGAGCTAGCCTCACAGTGCAGGCAGCTGCAGAGACT
TTGAGTGAAGTAGCCTAAAGGGATTTTAGAGAATCGTATTGAAATTGTTGTTGAGTGTCA
CTCAGAAAGAATCGTAATGGGTATATTGATAAATTAAAATTGGTATATTGAAATAAAAGT
TGAATATTATATAA

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756
>subunit 1 of 1, 750 aa, 1 stop
>MW: 84305, pI: 6.93, NX(S/T): 10
MWNLLEHTDSAVATARRPRWLCAAGALVLAGGFFLLGFLFGWFIKSSNEATNITPKHNMK AFL
DELAENIKKFLHNFTQIPLHLAGTEQNQQLAKQIQSQWKEFGLDSVELAHYDVLLSPNKTH
PNYISIINEDGNEIFNTSLFEPPPGYENVS DIVPPFSAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINCSGKIVIARYGVFRGNVKVNAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQRGNILNLNGAGDPLTPGYPANEYAYRRGIAEAVGLPSIPVPHPIGYYDAQKLLEKMG
GSAPPDSSWRGSLKV PNVNGPGFTGNFSTQVKV KMHIHSTNEVTRIYNVIGTLRGA VE PDRYV
ILGGHRDSWVFGGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFA SWDAEEFGLLGSTEW
AEENSRLLQERGVAYINADSSIEGNYTLRV DCTPLMYSILVHNLTKE LKSPDEGFEGKSLYES
WTKKSPSPEFGMPRISK LGSGND FEFV FQRLGIASGRARYTKNWE TNKFSGYPL YHSVYET
YELVEKFYDPMFKYH LTV A QVRGGMV FELANSIVLPFD CRDYAVV L RKYADKI YSI SMKHPQ
EMKTYSVSFDSLFS AVKNFTEIAS KFSERLQDFDKSNPIV LRM MN DQLMFLERA FIDPLGLP
DRPFYRHVIYAPSSH NKYAGESFP GIYDAL FDIES KVDP SKA WGEV KRQIY VAAFTVQAAA E
TLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713

009678544.101603